

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЗАХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**

***Методичні вказівки для виконання завдань самостійної роботи
з курсу
«Економетрика»***

Тернопіль – 2024

Рецензенти:

С. В. Мартинюк – к. ф.-м. н., доцент кафедри інформатики і методики її викладання Тернопільського національного педагогічного університету імені Володимира Гнатюка

О. С. Башуцька – к. е. н., доцент кафедри економічної кібернетики та інформатики Західноукраїнського національного університету

Затверджено на засіданні кафедри прикладної математики
протокол №1 від 26.08.2024.

Березька К. М., Мартинюк О. М., Пласконь С. А., Єрьоменко В. О., Руська Р. В., Маслій В. В. Завдання для самостійної роботи з курсу «Економетрика». Тернопіль: ЗУНУ, 2024. 68 с.

У посібнику наведено програму, приклади розв’язування задач та індивідуальні завдання для самостійної роботи з дисципліни «Економетрика» для студентів денної та заочної форм навчання.

Відповідальний за випуск: О. М. Мартинюк, кандидат фізико-математичних наук,
завідувач кафедри ПМ ЗУНУ

© Березька К., 2024

Зміст

Програма дисципліни «Економетрика»	4
Структура залікового кредиту дисципліни «Економетрика»	6
Тематика самостійної роботи	8
Завдання 1	8
Завдання 2	10
Завдання 3	11
Завдання 4	12
Завдання 5	12
Завдання 6	13
Завдання 7	14
Завдання 8	17
Завдання 9	17
Завдання 10	18
Зразки розв’язування самостійної роботи	19
ДОДАТКИ	58
Додаток 1. Алгоритм знаходження оберненої матриці в Excel	58
Додаток 2. Алгоритм множення матриць в Excel	57
Додаток 3. Порядок знаходження оцінок параметрів економетричної моделі з використанням функції LINEST	59
Додаток 4. Довідкові таблиці	61
СПИСОК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ	67

Програма дисципліни «Економетрика»

Змістовий модуль 1. Методологія побудови однофакторних економетричних моделей

Тема 1. Економетрія: основні поняття та визначення

Предмет та метод економетрії. Історичні відомості. Значення курсу та взаємозв'язок з іншими економічними дисциплінами.

Математична модель та основні етапи її побудови. Теоретичні основи математичного моделювання та класифікація моделей. Приклади моделей та методів, які носять і не носять характер економетричних досліджень. Підходи до вибору моделей та методів оцінки параметрів.

Тема 2. Формування та обробка даних

Типи даних. Методи збору даних у сучасній економіці. Основи обробки та підготовки даних для подальшого аналізу. Джерела даних. Методи перевірки та забезпечення точності отриманих даних. Типи шкал вимірювання змінних та їх використання в економетричних дослідженнях. Основні проблеми, що виникають під час обробки даних і методи їх вирішення.

Тема 3. Лінійна регресія із однією змінною

Лінійна регресійна модель та її використання у сфері управління персоналом. Знаходження статистичних оцінок параметрів моделі методом найменших квадратів (МНК). Коефіцієнти кореляції та детермінації. Побудова довірчих інтервалів. Перевірка нульових гіпотез. Адекватність регресійної моделі. Критерій Фішера. Прогнозування за моделлю лінійної регресії із однією змінною. Поняття гомо- і гетероскедастичності. Використання t -статистики в регресії для малої вибірки. Біноміальна регресійна модель. Приклади застосування моделей лінійної регресії.

Тема 4. Нелінійна регресія

Криві зростання. Зведення деяких нелінійних моделей до лінійних. Лінеаризація квадратичних функцій. Лінеаризація зворотних кривих зростання. Лінеаризація експоненційних функцій. Лінеаризація степеневих функцій. Приклади застосування нелінійних моделей.

Змістовий модуль 2. Методологія побудови багатофакторних економетричних моделей

Тема 5. Багатофакторна регресія

Модель множинної лінійної регресії у системі управління персоналом. МНК для багатофакторної регресійної моделі. Припущення методу найменших квадратів для причинного висновку в множинній регресії. Мультиколінераність: причини виникнення, методи оцінювання та способи усунення. Коефіцієнти парної, частинної та множинної кореляції, коефіцієнт детермінації. Оцінка якості параметрів моделі множинної регресії за методом найменших квадратів. Перевірка

моделі на адекватність. Перевірка гіпотез для двох і більше коефіцієнтів. *F*-критерій Фішера. Побудова довірчих інтервалів параметрів. Перевірка нульової гіпотези стосовно коефіцієнта множинної кореляції. Прогнозування за моделлю багатофакторної регресії. Модель множинної регресії з бінарними незалежними змінними. Загальний підхід до моделювання нелінійностей за допомогою множинної регресії. Неправильна специфікація функціональної форми регресійної моделі.

Змістовий модуль 3. Сучасні методи регресійного аналізу та прогнозування

Тема 6. Регресія з бінарною залежною змінною

Двійкові залежні змінні та модель лінійної ймовірності. *Probit* - і *Logit* - регресія. Порівняння моделей лінійної ймовірності, *Probit* та *Logit*. Оцінка та висновки у моделях *Logit* і *Probit*. Нелінійна оцінка методом найменших квадратів.

Тема 7. Регресійний аналіз на основі панельних даних

Панельні дані: характеристика та структура. Загальна регресійна модель панельних даних. Регресійна модель з фіксованими ефектами. Регресійна модель з випадковими ефектами.

Тема 8. Економетричний аналіз часових рядів

Застосування економетричних моделей для прогнозування. Методи виявлення тренду та сезонності в часових рядах. Стаціонарні часові ряди і їх характеристики в системі управління персоналом. Автокореляція: природа та наслідки. Методи визначення автокореляції. Автокореляційна функція. Методи згладжування динамічних рядів. Моделі Фур'є. Авторегресійні моделі динаміки та оцінювання їх параметрів. Прогнозування та побудова довірчих інтервалів.

Тема 9. Моделі розподіленого лагу

Поняття лагу і лагових змінних. Види лагових моделей. Автокореляція: природа та наслідки. Методи визначення автокореляції. Взаємна кореляційна функція. Лаги залежних і незалежних змінних. Методи їх оцінювання.

Тема 10. Прогнозування на основі великих даних

Що таке «великі дані»? Методи збору, обробки та підготовки великих обсягів даних для подальшого аналізу та прогнозування. Моделі прогнозування, придатні для великих обсягів даних. Оцінка точності та надійності прогнозів, здійснених на основі великих даних. Перехресна валідація.

Структура залікового кредиту дисципліни «Економетрика»
денна форма навчання

	Кількість годин				
	Лекції	Практичні заняття	ІРС	Тренінг	Контрольні заходи
Змістовий модуль 1. Методологія побудови однофакторних економетричних моделей					
Тема 1. Економетрія: основні поняття та визначення	2	1	1	2	Поточне опитування та тестування
Тема 2. Формування та обробка даних	2	1			Поточне опитування, тестування, задачі
Тема 3. Лінійна регресія із однією змінною	4	4			Поточне опитування, тестування, задачі
Тема 4. Нелінійна регресія	2	2			Поточне опитування, тестування, задачі
Змістовий модуль 2. Методологія побудови багатфакторних економетричних моделей					
Тема 5. Багатфакторна регресія	6	6	1	2	Поточне опитування, тестування, задачі
Змістовий модуль 3. Сучасні методи регресійного аналізу та прогнозування					
Тема 6. Регресія з бінарною залежною змінною	2	2	2	4	Поточне опитування, тестування, задачі
Тема 7. Регресійний аналіз на основі панельних даних	4	4			Поточне опитування, тестування, задачі
Тема 8. Економетричний аналіз часових рядів	4	4			Поточне опитування, тестування, задачі
Тема 9. Моделі розподіленого лагу	2	2			Поточне опитування, тестування, задачі
Тема 10. Прогнозування на основі великих даних	2	4			Поточне опитування, тестування, задачі
Всього	30	30	4	8	

заочна форма навчання

	Кількість годин	
	Лекції	Практичні заняття
Змістовий модуль 1. Методологія побудови однофакторних економетричних моделей		
Тема 1. Економетрія: основні поняття та визначення	1	-
Тема 2. Формування та обробка даних	1	1
Тема 3. Лінійна регресія із однією змінною	2	1
Тема 4. Нелінійна регресія	2	1
Змістовий модуль 2. Методологія побудови багатфакторних економетричних моделей		
Тема 5. Багатфакторна регресія	2	1
Змістовий модуль 3. Сучасні методи регресійного аналізу та прогнозування		
Тема 6. Регресія з бінарною залежною змінною	-	-

Тема 7. Регресійний аналіз на основі панельних даних	-	-
Тема 8. Економетричний аналіз часових рядів	-	-
Тема 9. Моделі розподіленого лагу	-	-
Тема 10. Прогнозування на основі великих даних	-	-
Всього	8	4

Тематика самостійної роботи

Завдання 1. Побудова лінійної моделі парної регресії.

Для десяти підприємств регіону за умовний деякий період відомі числові значення двох економічних показників: валова продукція Y (млн. грн.) і вартість основних виробничих фондів X (млн. грн.), (табл.1). Для дослідження характеристики впливу вартості основних виробничих фондів (X) на випуск валової продукції (Y) підприємства з допомогою економетричної моделі необхідно:

1. Обчислити статистичні оцінки параметрів лінійного рівняння регресії (за системою нормальних рівнянь та через відхилення від середніх);
2. Обчислити загальну, пояснену та непояснену дисперсії.
3. Для рівня значущості $\alpha = 0,05$ перевірити значущість коефіцієнтів регресії α_0 та α_1 ;
4. Знайти довірчі інтервали коефіцієнтів регресії з надійністю $\gamma = 0,95$;
5. Знайти вибіркові коефіцієнт детермінації, коефіцієнт кореляції;
6. Знайти та побудувати довірчий інтервал функції регресії з надійністю $p = 0,90$;
7. Виконати перевірку нульової гіпотези відносно коефіцієнта кореляції.
8. Перевірити значущість коефіцієнта детермінації з ймовірністю $p=0,95$.
9. Знайти прогнозне значення валової продукції, вартість основних виробничих фондів якої складає 8 млн. грн, а також із надійністю $\gamma = 0,95$ побудувати довірчий інтервал для цього прогнозного значення.

Числові параметри варіантів наведені в таблицях 2 та 3.

Таблиця 1

№ підприємства	Валовий випуск продукції, Y , млн. грн.	Вартість основних виробничих фондів, X , млн. грн.
1	$2,2 + a_1$	$1,3 + b_1$
2	$4,2 + a_2$	$2,2 + b_2$
3	$5,8 + a_3$	$3,4 + b_3$
4	$6,8 + a_4$	$2,6 + b_4$

5	$5,8 + a_5$	$3,2 + b_5$
6	$7,8 + a_6$	$4,5 + b_1$
7	$9,6 + a_7$	$5,0 + b_2$
8	$8,6 + a_8$	$6,2 + b_3$
9	$10,2 + a_9$	$7,3 + b_4$
10	$12,4 + a_{10}$	$8,2 + b_5$

Таблиця 2

Варіант, остання цифра № залікової	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
0	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	0.1	2.1
1	2.1	2.0	1.4	2.3	2.2	2.6	2.7	2.9	0.2	3.2
2	0.5	0.6	0.7	0.8	0.9	0.2	0.5	0.5	0.3	0.4
3	2.1	1.2	1.3	3.4	2.4	2.5	2.6	2.7	0.4	3.1
4	1.5	2.1	1.2	5.1	1.2	2.0	1.8	3.2	0.1	3.1
5	1.6	2.2	1.3	5.2	1.3	2.1	1.9	3.4	0.2	3.2
6	1.7	2.3	1.4	5.3	1.4	2.2	1.7	3.5	0.3	3.3
7	1.8	2.4	1.6	5.4	1.5	2.3	1.4	4.1	0.4	3.4
8	1.9	2.6	1.7	5.2	1.6	2.4	1.5	4.2	0.4	3.6
9	2.0	2.7	1.8	5.0	1.7	2.6	1.2	4.3	0.2	3.7

Таблиця 3

Номер варіанту вибирати за порядковим номером студента в списку групи

Варіант	b_1	b_2	b_3	b_4	b_5	Варіант	b_1	b_2	b_3	b_4	b_5
1	0.5	0.6	0.8	0.9	0.9	24	2.1	1.3	4.1	2.1	0.3
2	1.1	1.2	1.1	1.2	1.2	25	0.5	0.6	1.2	1.4	0.6
3	1.3	1.1	1.4	1.5	1.2	26	0.5	0.4	0.7	1.0	1.0
4	1.6	1.2	1.3	1.4	1.5	27	1.2	1.3	1.4	2.0	1.5
5	1.6	1.7	1.8	1.4	1.5	28	2.0	2.1	0.3	0.3	1.1
6	0.3	0.3	0.2	0.2	0.3	29	1.2	1.6	1.7	0.3	0.4
7	1.0	1.2	1.2	1.3	1.3	30	1.5	1.4	0.8	0.3	0.2
8	1.3	1.5	1.7	1.8	0.2	31	0.8	0.2	0.1	2.1	1.3
9	0.3	1.2	1.4	0.2	0.1	32	0.9	0.5	1.9	1.3	1.4
10	1.8	1.1	0.3	0.5	0.7	33	1.7	1.7	1.7	1.8	2.0
11	0.7	0.8	2.1	2.3	1.5	34	0.9	0.8	1.3	1.2	2.3
12	1.9	1.3	0.2	1.3	1.6	35	2.4	1.4	1.5	0.3	0.5
13	2.2	2.4	1.5	1.7	0.2	36	1.5	1.8	0.6	0.7	0.3
14	0.3	0.4	0.5	0.7	0.8	37	1.4	1.7	1.9	2.3	1.8
15	0.4	1.3	1.4	1.7	1.2	38	2.1	0.4	0.6	1.5	1.2
16	2.3	1.0	1.3	0.5	0.6	39	1.9	1.3	1.1	2.1	0.8
17	0.7	0.3	1.2	1.4	1.1	40	2.4	2.1	0.9	1.8	1.9
18	1.1	1.4	1.3	1.0	2.3	41	2.1	0.8	0.7	2.3	1.2
19	2.2	2.3	1.5	0.2	0.6	42	2.3	0.3	0.4	0.9	1.5

20	1.6	1.7	2.1	0.5	1.6	43	1.6	1.7	0.9	2.2	0.7
21	1.7	1.2	1.8	1.9	1.0	44	0.7	0.8	1.3	1.4	1.4
22	2.3	2.1	1.8	2.0	1.2	45	0.9	0.7	1.5	1.6	0.5
23	2.4	2.2	0.9	0.7	0.8	46	0.4	0.3	2.1	2.4	1.5

Завдання 2. Нелінійні економетричні моделі

- Використовуючи вибірккові дані завдання 1, підібрати криву, яка найповніше описує тенденцію з допомогою Excel (ЛИНИЯ ТРЕНДА). Для кращого підбору лінії тренду пробувати декілька варіантів кривих: лінійна; поліноміальна (ступінь 2); поліноміальна (ступінь 3); логарифмічна; степенева; експоненційна.
- Побудувати залежність виду згідно свого варіанту (див. табл. 4), використавши МНК за системою нормальних рівнянь.

Таблиця 4

Варіант (за останньою цифрою залікової книжки)	Вид залежності
0	$y = \alpha_0 + \alpha_1 \frac{1}{x}$
1	$y = \alpha_0 + \alpha_1 \sqrt{x}$
2	$y = \alpha_0 + \alpha_1 x^3$
3	$y = \alpha_0 + \alpha_1 \sqrt{\frac{1}{x}}$
4	$y = \alpha_0 + \alpha_1 x^2$
5	$y = \alpha_0 + \alpha_1 \ln x$
6	$y = \alpha_0 \alpha_1^x$
7	$y = \alpha_0 x^{\alpha_1}$
8	$y = \alpha_0 + \alpha_1 \sqrt[3]{x}$
9	$y = e^{\alpha_0 + \alpha_1 x}$

Завдання 3. Багатофакторна регресія.

Знаходження оцінок параметрів регресійної моделі методом найменших квадратів з застосуванням системи нормальних рівнянь

На основі даних *таблиці 5*:

- 1) оцінити параметри лінійної регресійної моделі, яка описує залежність валової продукції від вартості основних виробничих фондів та затрат робочого часу;
 - 2) обчислити коефіцієнти парної та частинної кореляції;
 - 3) знайти коефіцієнти множинної детермінації та кореляції.
- Зробити економічний аналіз отриманих розрахунків.

Числові значення даних для варіантів наведені в таблицях 5 та 2 і 3.

Таблиця 5

№ п/п	Валова продукція, Y , млн. грн.	Основні виробничі фонди, X_1 , млн. грн.	Затрати робочого часу, X_2 , тис. люд.-год.
1	$2,2 + a_1$	$1,3 + b_1$	$2,3 + a_1$
2	$4,2 + a_2$	$2,2 + b_2$	$3,4 + a_2$
3	$5,8 + a_3$	$3,4 + b_3$	$2,2 + a_3$
4	$6,8 + a_4$	$2,6 + b_4$	$3,4 + a_1$
5	$5,8 + a_5$	$3,2 + b_5$	$4,2 + a_2$
6	$7,8 + a_6$	$4,5 + b_1$	$4,3 + a_3$
7	$9,6 + a_7$	$5,0 + b_2$	$5,4 + a_1$
8	$8,6 + a_8$	$6,2 + b_3$	$5,6 + a_2$
9	$10,2 + a_9$	$7,3 + b_4$	$5,7 + a_3$
10	$12,4 + a_{10}$	$8,2 + b_5$	$6,2 + a_2$

Завдання 4. Багатофакторна регресія.

Знаходження оцінок параметрів регресійної моделі методом найменших квадратів з застосуванням матричної форми запису
Використовуючи вибірккові дані попереднього завдання, знайти:

1. Вектор оцінок $\mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$

2. Матрицю дисперсій оцінок.

3. Інтервали довір'я ($p=0,9$) параметрів α_0 , α_1 та α_2 .

Завдання 5. Мультиколінеарність в множинних регресійних моделях

Економічний показник середньомісячна заробітна плата (y) залежить від продуктивності праці (x_1), фондомісткості (x_2) і коефіцієнта плинності робочої сили (x_3). На основі статистичних даних за 10 років необхідно оцінити наявність загальної мультиколінеарності. У випадку її присутності, виявити пари факторів між якими існує мультиколінеарність, один із факторів виключити з розгляду таких пар. Дослідження провести за методом Феррара-Глобера.

Вихідні дані наводяться в табл. 6 та табл. 7

Таблиця 6

Номер цеху	Продуктивність праці, людино-дні	Фондомісткість, млн. грн.	Коефіцієнт плинності робочої сили, %
1	$32 + a_4$	$0.89 + a_9$	$19.5 + a_1$
2	$29 + a_5$	$0.43 + a_9$	$15.6 + a_2$
3	$30 + a_6$	$0.70 + a_9$	$13.5 + a_3$
4	$31 + a_7$	$0.61 + a_9$	$9.5 + a_1$
5	$25 + a_8$	$0.51 + a_9$	$23.5 + a_2$
6	$34 + a_4$	$0.71 + a_9$	$12.5 + a_3$
7	$29 + a_5$	$0.65 + a_9$	$17.5 + a_1$
8	$24 + a_6$	$0.43 + a_9$	$14.5 + a_2$
9	$20 + a_7$	$0.33 + a_9$	$14.5 + a_3$
10	$33 + a_8$	$0.92 + a_9$	$75 + a_1$

Таблиця 7

Варіант остання цифра № залікової	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
0	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	0.01	2.1
1	2.1	2.0	1.4	2.3	2.2	2.6	2.7	2.9	0.02	3.2

2	0.5	0.6	0.7	0.8	0.9	0.12	0.15	0.25	0.03	0.4
3	2.1	1.2	1.3	3.4	2.4	2.5	2.6	2.7	0.04	3.14
4	1.5	2.1	1.2	5.1	1.2	2.0	1.8	3.2	0.01	3.1
5	1.6	2.2	1.3	5.2	1.3	2.1	1.9	3.4	0.02	3.2
6	1.7	2.3	1.4	5.3	1.4	2.2	1.7	3.5	0.03	3.3
7	1.8	2.4	1.6	5.4	1.5	2.3	1.4	4.1	0.04	3.4
8	1.9	2.6	1.7	5.2	1.6	2.4	1.5	4.2	0.04	3.6
9	2.0	2.7	1.8	5.0	1.7	2.6	1.2	4.3	0.02	3.7

Завдання 6. Гетероскедастичність в лінійних моделях регресії

Побудувати лінійну модель парної регресії залежності величини доходу від рівня заощаджень. Перед побудовою залежності перевірити наявність явища гетероскедастичності з допомогою тестів: рангової кореляції Спірмена, Голдфелда-Квондта, Глейзера. Вихідні дані наведені табл. 8, а також у табл. 7 та 9.

Таблиця 8

Місяць	Дохід, умовних одиниць	Заощадження, умовних одиниць
1	$10.8 + b_1$	$2.36 + a_9$
2	$11.4 + b_2$	$2.20 + a_9$
3	$12.0 + b_3$	$2.08 + a_9$
4	$12.6 + b_4$	$2.20 + a_9$
5	$13.0 + b_5$	$2.10 + a_9$
6	$13.9 + b_1$	$2.12 + a_9$
7	$14.7 + b_2$	$2.41 + a_9$
8	$15.5 + b_3$	$2.50 + a_9$
9	$16.3 + b_4$	$2.43 + a_9$
10	$17.5 + b_5$	$2.59 + a_9$
11	$18.7 + b_5$	$2.9 + a_9$
12	$19.7 + b_4$	$2.95 + a_9$
13	$20.6 + b_3$	$2.82 + a_9$
14	$21.7 + b_2$	$3.04 + a_9$
15	$23.2 + b_1$	$3.53 + a_9$
16	$24.2 + b_1$	$3.44 + a_9$
17	$25.9 + b_2$	$3.75 + a_9$
18	$27.2 + b_3$	$3.99 + a_9$

Таблиця 9

Варіант, № у списку	b_1	b_2	b_3	b_4	b_5	Варіант, № у списку	b_1	b_2	b_3	b_4	b_5
1	0.5	0.6	0.8	0.9	0.9	26	2.3	1.0	1.3	0.5	0.6
2	1.1	1.2	1.1	1.2	1.2	27	0.7	0.3	1.2	1.4	1.1
3	1.3	1.1	1.4	1.5	1.2	28	1.1	1.4	1.3	1.0	2.3
4	1.6	1.2	1.3	1.4	1.5	29	2.2	2.3	1.5	0.2	0.6
5	1.6	1.7	1.8	1.4	1.5	30	1.6	1.7	2.1	0.5	1.6
6	0.3	0.3	0.2	0.2	0.3	31	1.7	1.2	1.8	1.9	2.0
7	1.0	1.2	1.1	1.3	1.3	32	2.3	2.1	1.8	2.0	1.2
8	1.3	1.5	1.7	1.8	0.2	33	2.4	2.2	0.9	0.7	0.8
9	0.3	1.2	1.4	0.2	0.1	34	1.8	1.9	2.1	0.5	0.9
10	1.8	1.1	0.3	0.5	0.7	35	2.3	2.1	2.4	2.0	0.8
11	2.1	1.3	4.1	2.1	0.3	36	0.9	0.8	1.3	1.2	2.3
12	0.5	0.6	1.2	1.4	0.6	37	2.4	1.4	1.5	0.3	0.5
13	0.5	0.4	0.7	1.0	1.0	38	1.5	1.8	0.6	0.7	0.3
14	1.2	1.3	1.4	2.0	2.5	39	1.4	1.7	1.9	2.3	1.8
15	2.0	2.1	0.3	0.3	1.1	40	2.1	0.4	0.6	1.5	1.2
16	1.2	1.6	1.7	0.3	0.4	41	1.9	1.3	1.1	2.1	0.8
17	1.5	1.4	0.8	0.3	0.2	42	2.4	2.1	0.9	1.8	1.9
18	0.8	0.2	0.1	2.1	2.3	43	2.1	0.8	0.7	2.3	1.2
19	0.9	0.5	1.9	1.3	1.4	44	2.3	0.3	0.4	0.9	1.5
20	1.7	1.7	1.7	1.8	2.0	45	1.6	1.7	0.9	2.2	0.7
21	0.7	0.8	2.1	2.3	1.5	46	0.7	0.8	1.3	1.4	2.4
22	1.9	1.3	0.2	1.3	1.6	47	0.9	0.7	1.5	1.6	0.5
23	2.2	2.4	1.5	1.7	0.2	48	0.4	0.3	2.1	2.4	1.5
24	0.3	0.4	0.5	0.7	0.8	49	1.3	1.1	0.4	1.5	1.7
25	0.4	1.3	1.4	1.7	1.2	50	1.9	2.4	1.8	0.7	0.3

Завдання 7. Автокореляція в лінійних моделях регресії.

Побудувати лінійну модель множинної регресії залежності прибутку підприємства від вартості основних виробничих фондів, затрат праці та собівартості продукції, попередньо дослідивши наявність автокореляції.

Числові параметри варіантів наведені в табл. 10, а також в табл. 7 та 9.

Таблиця 10

№ підприємства	Прибуток, Y , млн. грн.	Вартість основних виробничих фондів, x_1 , млн. грн.	Затрати праці, x_2 , тис. люд.днів	Собівартість одиниці продукції, x_3 , грн
1	$10.6 + a_3$	$20.4 + b_1$	$100.4 + a_6$	$10.3 + a_1$
2	$20.4 + a_4$	$30.5 + b_2$	$110.2 + a_7$	$20.3 + a_2$
3	$22.4 + a_5$	$25.6 + b_3$	$105.3 + a_8$	$24.5 + a_3$
4	$30.6 + a_6$	$40.3 + b_4$	$120.3 + a_9$	$28.3 + a_1$
5	$35.7 + a_7$	$60.5 + b_5$	$140.3 + a_{10}$	$29.3 + a_2$
6	$40.3 + a_4$	$65.3 + b_5$	$145.4 + a_1$	$35.2 + a_3$
7	$51.3 + a_4$	$70.8 + b_4$	$155.2 + a_2$	$38.4 + a_1$
8	$55.5 + a_5$	$85.1 + b_3$	$170.3 + a_3$	$48.2 + a_2$
9	$68.3 + a_6$	$90.2 + b_2$	$175.4 + a_4$	$50.4 + a_3$
10	$72.2 + a_7$	$100.2 + b_1$	$185.3 + a_5$	$58.2 + a_1$
11	$84.8 + a_3$	$120.4 + b_2$	$189.4 + a_1$	$55.3 + a_2$
12	$90.3 + a_5$	$125.6 + b_3$	$192.3 + a_6$	$60.3 + a_2$
13	$98.8 + a_6$	$132.4 + b_4$	$190.4 + a_5$	$69.4 + a_4$
14	$105.2 + a_7$	$136.7 + b_5$	$199.5 + a_3$	$70.3 + a_2$
15	$114.6 + a_3$	$145.2 + b_1$	$210.6 + a_9$	$75.8 + a_1$

Завдання 8. Регресійний аналіз на основі панельних даних.

Задача. Використовуючи статистичні дані (табл. 11), де Y – прибуток (млн. грн., результативна або залежна змінна); X_1 – обсяг інвестицій (млн. грн., пояснююча змінна); X_2 – кількість працівників (тис.ос., пояснююча змінна). проаналізувати вплив інвестицій та кількості працівників на прибуток, використовуючи панельні дані за 2020-2023 роки для компанії, що об'єднує три підприємства.

Числові параметри варіантів наведені в табл. 12 та табл. 13.

Таблиця 11

Рік	Підприємство	Y	X_1	X_2
2020	1	$21 + a_1$	$5 + a_1$	$11 + b_1$
2020	2	$25 + a_2$	$6 + a_2$	$12 + b_2$
2020	3	$30 + a_3$	$8 + a_3$	$14 + b_3$
2021	1	$22 + a_4$	$6 + a_4$	$11 + b_4$
2021	2	$27 + a_5$	$7 + a_5$	$13 + b_5$
2021	3	$32 + a_6$	$9 + a_6$	$15 + b_1$
2022	1	$24 + a_7$	$7 + a_7$	$1 + b_2$
2022	2	$28 + a_8$	$8 + a_8$	$14 + b_3$
2022	3	$34 + a_9$	$10 + a_9$	$16 + b_4$
2023	1	$26 + a_{10}$	$8 + a_{10}$	$13 + b_5$
2023	2	$30 + a_2$	$9 + a_2$	$15 + b_1$
2023	3	$36 + a_{10}$	$11 + a_{10}$	$17 + b_2$

Таблиця 12

Варіант, остання цифра № залікової	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
0	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	0.1	2.1
1	2.1	2.0	1.4	2.3	2.2	2.6	2.7	2.9	0.2	3.2
2	0.5	0.6	0.7	0.8	0.9	0.2	0.5	0.5	0.3	0.4
3	2.1	1.2	1.3	3.4	2.4	2.5	2.6	2.7	0.4	3.1
4	1.5	2.1	1.2	5.1	1.2	2.0	1.8	3.2	0.1	3.1
5	1.6	2.2	1.3	5.2	1.3	2.1	1.9	3.4	0.2	3.2
6	1.7	2.3	1.4	5.3	1.4	2.2	1.7	3.5	0.3	3.3
7	1.8	2.4	1.6	5.4	1.5	2.3	1.4	4.1	0.4	3.4
8	1.9	2.6	1.7	5.2	1.6	2.4	1.5	4.2	0.4	3.6
9	2.0	2.7	1.8	5.0	1.7	2.6	1.2	4.3	0.2	3.7

Таблиця 13

Номер варіанту вибирати за порядковим номером студента в списку групи

Варіант	b_1	b_2	b_3	b_4	b_5	Варіант	b_1	b_2	b_3	b_4	b_5
1	0.5	0.6	0.8	0.9	0.9	24	2.1	1.3	4.1	2.1	0.3
2	1.1	1.2	1.1	1.2	1.2	25	0.5	0.6	1.2	1.4	0.6
3	1.3	1.1	1.4	1.5	1.2	26	0.5	0.4	0.7	1.0	1.0
4	1.6	1.2	1.3	1.4	1.5	27	1.2	1.3	1.4	2.0	1.5
5	1.6	1.7	1.8	1.4	1.5	28	2.0	2.1	0.3	0.3	1.1
6	0.3	0.3	0.2	0.2	0.3	29	1.2	1.6	1.7	0.3	0.4
7	1.0	1.2	1.2	1.3	1.3	30	1.5	1.4	0.8	0.3	0.2
8	1.3	1.5	1.7	1.8	0.2	31	0.8	0.2	0.1	2.1	1.3
9	0.3	1.2	1.4	0.2	0.1	32	0.9	0.5	1.9	1.3	1.4
10	1.8	1.1	0.3	0.5	0.7	33	1.7	1.7	1.7	1.8	2.0
11	0.7	0.8	2.1	2.3	1.5	34	0.9	0.8	1.3	1.2	2.3
12	1.9	1.3	0.2	1.3	1.6	35	2.4	1.4	1.5	0.3	0.5
13	2.2	2.4	1.5	1.7	0.2	36	1.5	1.8	0.6	0.7	0.3
14	0.3	0.4	0.5	0.7	0.8	37	1.4	1.7	1.9	2.3	1.8
15	0.4	1.3	1.4	1.7	1.2	38	2.1	0.4	0.6	1.5	1.2
16	2.3	1.0	1.3	0.5	0.6	39	1.9	1.3	1.1	2.1	0.8
17	0.7	0.3	1.2	1.4	1.1	40	2.4	2.1	0.9	1.8	1.9
18	1.1	1.4	1.3	1.0	2.3	41	2.1	0.8	0.7	2.3	1.2
19	2.2	2.3	1.5	0.2	0.6	42	2.3	0.3	0.4	0.9	1.5
20	1.6	1.7	2.1	0.5	1.6	43	1.6	1.7	0.9	2.2	0.7
21	1.7	1.2	1.8	1.9	1.0	44	0.7	0.8	1.3	1.4	1.4
22	2.3	2.1	1.8	2.0	1.2	45	0.9	0.7	1.5	1.6	0.5
23	2.4	2.2	0.9	0.7	0.8	46	0.4	0.3	2.1	2.4	1.5

Завдання 9. Економетричний аналіз часових рядів.

Задача. Дослідити та побудувати тренд на основі даних прямих іноземних інвестицій в економіку країни (табл. 15), якщо основна тенденція динаміки описується прямою лінією або поліноміальною функцією. Розрахувати прогнозний показник прямих іноземних інвестицій на наступний рік.

Параметри для обчислення вибираються з таблиці 12.

Таблиця 15

Рік	1	2	3	4	5	6
Обсяг іноземних	750,0+	1399,6+	2041,9+	2794,9+	3287,4+	3918,5+

інвестицій, млн. дол.	a_1	a_2	a_3	a_4	a_5	a_6
Рік	7	8	9	10	11	12
Обсяг іноземних інвестицій, млн. дол.	4406,2+ a_7	5339,0+ a_8	6657,6+ a_9	8353,9+ a_{10}	16375,2+ a_1	21186,0+ a_2

Завдання 10. Моделі із лаговими змінними

Задача. Розрахувати значення коефіцієнтів взаємної кореляції та побудувати графік кореляційної функції для наступних статистичних даних залежності витрат на харчування (y_t , грош. од.) від рівня доходів (x_t , грош. од.) при максимальному значенні лага часу, рівному шести місячним періодам.

Параметри для чисельних даних вибрати з таблиць 12 та 13.

Рік	y_t	x_t
1	1400+ a_1	1150+ b_1
2	1560+ a_2	1270+ b_2
3	1720+ a_3	1390+ b_3
4	1500+ a_4	1400+ b_4
5	1250+ a_5	1200+ b_5
6	1320+ a_6	1350+ b_1
7	1429+ a_7	1400+ b_2
8	1560+ a_8	1410+ b_3
9	1600+ a_9	1450+ b_4
10	1713+ a_{10}	1500+ b_5
11	1800+ a_1	1400+ b_1
12	1780+ a_2	1530+ b_2
13	1850+ a_3	1540+ b_3
14	1900+ a_4	1500+ b_4
15	2100+ a_5	1600+ b_5

Зразки розв'язування самостійної роботи

Завдання 1. Лінійна регресія із однією змінною

Для десяти підприємств регіону за умовний деякий період відомі числові значення двох економічних показників: валова продукція y і вартість основних виробничих фондів x , (табл.12). Для дослідження характеристики впливу вартості основних виробничих фондів (x) на випуск валової продукції (y) підприємства з допомогою економетричної моделі необхідно:

1. Зробити специфікацію моделі.
2. Знайти оцінки параметрів лінійної моделі з допомогою МНК (за системою нормальних рівнянь та через відхилення від середніх).
3. Побудувати оціночну пряму.
4. Обчислити загальну, пояснюючу та непояснюючу дисперсію.
5. Знайти значення вибіркових коефіцієнтів детермінації і кореляції та показників $MARE$ і MPE .
6. Побудувати довірчі інтервали параметрів α_0 та α_1 економетричної моделі з надійністю 0,95.
7. Побудувати довірчий інтервал рівняння економетричної моделі з надійністю 0,95.
8. Виконати перевірки нульових гіпотез стосовно коефіцієнта кореляції і параметрів α_0 та α_1 економетричної моделі для рівня значущості $\alpha=0,05$.
9. Перевірити адекватність побудованої економетричної моделі.

Таблиця 12

№ підприємства	Валовий випуск продукції, млн.грн., y_i	Основні виробничі фонди, млн.грн., x_i
1	1,9	1,8
2	2,2	2,1
3	3	2,4
4	3,1	2,3
5	3,9	2,6
6	4,6	2,9
7	4,8	2,7
8	5,8	3,1
9	7,3	3,5
10	8,7	4,1

♦Розв'язування.

1. Побудуємо діаграму розсіювання залежності валового випуску продукції (y) від вартості основних виробничих фондів підприємства (x), відклавши на координатній площині точки з координатами $(x_i; y_i)$ з таблиці 12:

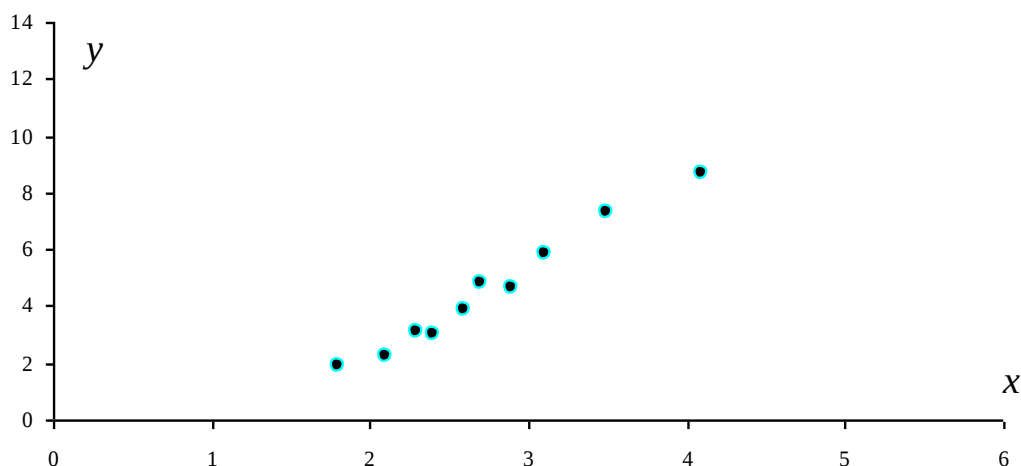


Рис.1. Діаграма розсіювання

Розміщення точок на діаграмі розсіювання дає можливість зробити припущення про існування лінійної форми зв'язку у вигляді функції:

$$\hat{y} = a_0 + a_1 x,$$

де $\hat{y}$ – розрахунковий обсяг випуску валової продукції, млн. грн.; x – вартість основних виробничих фондів, млн. грн.

2. Для спрощення розрахунків при знаходженні оцінок a_0 та a_1 параметрів економетричної моделі побудуємо таблицю:

Таблиця 13

№ п/п	y_i	x_i	x_i^2	$x_i \cdot y_i$	Δx_i	$(\Delta x_i)^2$	Δy_i	$\Delta x_i \cdot \Delta y_i$
1	1,9	1,8	3,24	3,42	-0,95	0,903	-2,63	2,499
2	2,2	2,1	4,41	4,62	-0,65	0,423	-2,33	1,515
3	3	2,4	5,76	7,2	-0,35	0,123	-1,53	0,536
4	3,1	2,3	5,29	7,13	-0,45	0,203	-1,43	0,644
5	3,9	2,6	6,76	10,14	-0,15	0,023	-0,63	0,094
6	4,6	2,9	8,41	13,34	0,15	0,023	0,07	0,011
7	4,8	2,7	7,29	12,96	-0,05	0,002	0,27	-0,014
8	5,8	3,1	9,61	17,98	0,35	0,123	1,27	0,445
9	7,3	3,5	12,25	25,55	0,75	0,563	2,77	2,078
10	8,7	4,1	16,81	35,67	1,35	1,823	4,17	5,63
Сума	45,3	27,5	79,83	138,01	0	4,205	0	13,435

Δx_i – це відхилення кожного значення x_i від середнього $\bar{x}$;

Δy_i – це відхилення кожного значення y_i від середнього $\bar{y}$;

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{27,5}{10} = 2,75; \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{45,3}{10} = 4,53.$$

$$\Delta x_1 = x_1 - \bar{x} = 1,8 - 2,75 = -0,95;$$

$$\Delta x_2 = x_2 - \bar{x} = 2,1 - 2,75 = -0,65;$$

.....

$$\Delta x_{10} = x_{10} - \bar{x} = 4,1 - 2,75 = 1,35.$$

Запишемо систему нормальних рівнянь:

$$\begin{cases} na_0 + a_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \\ a_0 \sum_{i=1}^n x_i + a_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \end{cases} \Rightarrow \begin{cases} 10a_0 + 27,5a_1 = 45,3 \\ 27,5a_0 + 79,83a_1 = 138,01 \end{cases}$$

Розв'яжемо її:

$$\begin{cases} a_0 = -2,75a_1 + 4,53 \\ 27,5(-2,75a_1 + 4,53) + 79,83a_1 = 138,01 \end{cases} \Rightarrow$$

$$\begin{cases} a_0 = -2,75a_1 + 4,53 \\ -75,625a_1 + 124,575 + 79,83a_1 = 138,01 \end{cases} \Rightarrow$$

$$\begin{cases} a_0 = -2,75a_1 + 4,53 \\ 4,205a_1 + 124,575 = 138,01 \end{cases} \Rightarrow \begin{cases} a_0 = -2,75a_1 + 4,53 \\ 4,205a_1 = 13,435 \end{cases} \Rightarrow$$

$$\begin{cases} a_0 = -2,75 \cdot 3,195 + 4,53 \\ a_1 = 3,195 \end{cases} \Rightarrow \begin{cases} a_0 = -4,26 \\ a_1 = 3,195 \end{cases}$$

Знайдемо ці ж оцінки за формулами відхилень від середніх:

$$a_1 = \frac{\sum_{i=1}^n (\Delta x_i \Delta y_i)}{\sum_{i=1}^n (\Delta x_i)^2} = \frac{13,435}{4,205} = 3,195;$$

$$a_0 = \bar{y} - a_1 \bar{x} = 4,53 - 3,195 \cdot 2,75 = -4,26.$$

Отже, нами отримано оціночне рівняння економетричної моделі
 $\hat{y} = -4,26 + 3,195x$.

3. Побудуємо оціночну пряму $\hat{y} = -4,26 + 3,195x$.

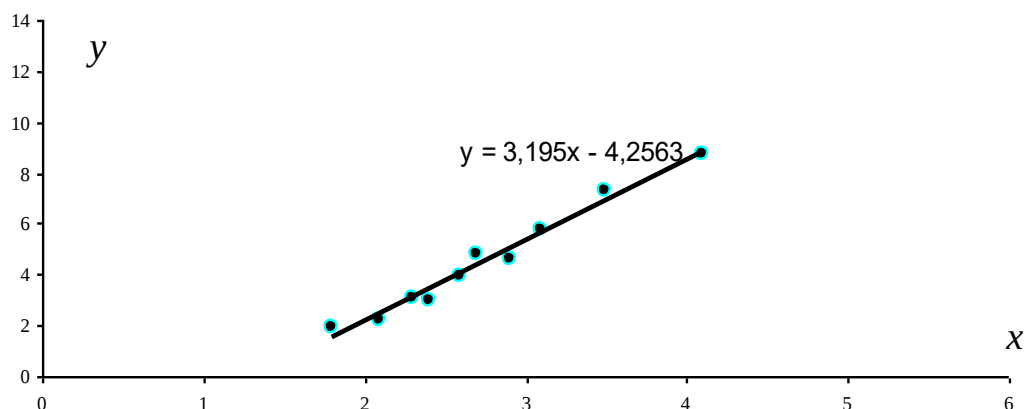


Рис.2. Діаграма розсіювання та оціночна пряма

4. Для знаходження дисперсій використаємо наступні формули:

$$\sigma_{\text{заг.}}^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}; \quad \sigma_{\text{поясн.}}^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{n}; \quad \sigma_{\text{непоясн.}}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}.$$

Для спрощення підрахунків побудуємо таблицю, використавши середні значення змінних: $\bar{x} = 2,75$; $\bar{y} = 4,53$.

Таблиця 14

№ п/п	y_i	x_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$\hat{y}_i$	$\hat{y}_i - \bar{y}$	$(\hat{y}_i - \bar{y})^2$	$\hat{y}_i - y_i$	$(y_i - \hat{y}_i)^2$
1	1,9	1,8	-2,63	6,92	1,49	-3,04	9,21	-0,41	0,164
2	2,2	2,1	-2,33	5,43	2,45	-2,08	4,31	0,25	0,064
3	3	2,4	-1,53	2,34	3,41	-1,12	1,25	0,41	0,17
4	3,1	2,3	-1,43	2,04	3,09	-1,44	2,07	-0,01	0,0001
5	3,9	2,6	-0,63	0,4	4,05	-0,48	0,23	0,15	0,023
6	4,6	2,9	0,07	0	5,01	0,48	0,23	0,41	0,167
7	4,8	2,7	0,27	0,07	4,37	-0,16	0,03	-0,43	0,185
8	5,8	3,1	1,27	1,61	5,65	1,12	1,25	-0,15	0,023
9	7,3	3,5	2,77	7,67	6,93	2,4	5,74	-0,37	0,14
10	8,7	4,1	4,17	17,39	8,84	4,31	18,6	0,14	0,021
Сума	45,3	27,5	0	43,88	45,3	0	42,92	0	0,956

Отже, маємо:

$$\sigma_{\text{заг.}}^2 = \frac{43,88}{10} = 4,388; \quad \sigma_{\text{поясн.}}^2 = \frac{42,92}{10} = 4,292; \quad \sigma_{\text{непоясн.}}^2 = \frac{0,956}{10} = 0,096.$$

5. Знайдемо значення вибірових коефіцієнтів детермінації та кореляції.

Для обчислення коефіцієнта детермінації використовуємо формулу:

$$d = \frac{\sigma_{\text{поясн.}}^2}{\sigma_{\text{заг.}}^2} = \frac{4,292}{4,388} = 0,978,$$

а це означає, що 97,8 % загальної дисперсії пояснюється оціночною прямою, на долю неврахованих факторів припадає 2,2 %.

Коефіцієнт кореляції знайдемо за формулою:

$$r = \pm\sqrt{d} = \pm\sqrt{0,978} = 0,99.$$

Знак коефіцієнта кореляції визначається знаком кутового коефіцієнта оціночного рівняння a_1 (в нашому випадку він додатний). Отримане значення коефіцієнта кореляції вказує на ступінь тісноти лінійного зв'язку між змінними. Значення коефіцієнта кореляції рівне 0,99 (дуже близьке до одиниці), а це значить, що лінійна форма зв'язку між змінними y та x вибрана вірно і цей зв'язок тісний.

Обчислимо значення абсолютної середньої відсоткової помилки MAPE за формулою:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \cdot 100\%.$$

Для цього використаємо розрахунки таблиці 3:

$$\begin{aligned} \sum_{i=1}^{10} \left| \frac{\hat{y}_i - y_i}{y_i} \right| &= \left| \frac{-0,41}{1,9} \right| + \left| \frac{0,25}{2,2} \right| + \left| \frac{0,41}{3} \right| + \left| \frac{-0,01}{3,1} \right| + \left| \frac{0,15}{3,9} \right| + \left| \frac{0,41}{4,6} \right| + \left| \frac{-0,43}{4,8} \right| + \\ &+ \left| \frac{-0,15}{5,8} \right| + \left| \frac{-0,37}{7,3} \right| + \left| \frac{0,14}{8,7} \right| = 0,213 + 0,115 + 0,137 + 0,003 + 0,039 + \\ &+ 0,089 + 0,09 + 0,026 + 0,051 + 0,017 = 0,78. \end{aligned}$$

Тоді $MAPE = \frac{1}{10} \cdot 0,78 \cdot 100\% = 7,8\% < 10\%$, що відповідає високій точності прогнозу за отриманою економетричною моделлю $\hat{y} = -4,26 + 3,195x$.

Середню відсоткову помилку MPE знайдемо за формулою:

$$MPE = \frac{1}{n} \sum_{i=1}^n \frac{\hat{y}_i - y_i}{y_i} \cdot 100\%,$$

Обчислимо

$$\begin{aligned} \sum_{i=1}^{10} \frac{\hat{y}_i - y_i}{y_i} &= -0,213 + 0,115 + 0,137 - 0,003 + 0,039 + \\ &+ 0,089 - 0,09 - 0,026 - 0,051 + 0,017 = 0,014. \end{aligned}$$

Отримали

$$MPE = \frac{1}{10} \cdot 0,014 \cdot 100\% = 0,14\% < 5\%, \text{ що свідчить про високу якість}$$

економетричної моделі.

6. Довірчі інтервали з надійністю γ для невідомих параметрів a_0 та a_1 мають вигляд:

$$a_m - t_m(\gamma, k)S_{a_m} < \alpha_m < a_m + t_m(\gamma, k)S_{a_m},$$

де $m = 0; 1$, $t_m = t_m(\gamma, k)$ – корінь рівняння $P(|t_m| < t) = \gamma$, t_m – випадкова величина, розподілена за законом Ст'юдента.

Для надійності $\gamma = 0,95$ і числа ступенів вільності $k = n - 2 = 8$ за таблицею значень функції Ст'юдента знаходимо $t_0(0,95; 8) = t_1(0,95; 8) = 2,306$. Потім обчислимо незміщену точкову оцінку σ_u^2 непоясненої дисперсії за формулою:

$$\sigma_u^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \frac{1}{10-2} \cdot 0,956 = 0,1195.$$

Значення $\sigma_{a_0}^2$, $\sigma_{a_1}^2$ та σ_{a_0} , σ_{a_1} знаходимо за формулами:

$$\sigma_{a_0}^2 = \frac{\sigma_u^2 \cdot \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}, \quad \sigma_{a_0} = \sqrt{\sigma_{a_0}^2}; \quad \sigma_{a_1}^2 = \frac{\sigma_u^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad \sigma_{a_1} = \sqrt{\sigma_{a_1}^2}.$$

$$\sigma_{a_0}^2 = \frac{0,1195 \cdot 79,83}{10 \cdot 4,205} = 0,227, \quad \sigma_{a_0} = \sqrt{0,227} = 0,476;$$

$$\sigma_{a_1}^2 = \frac{0,1195}{4,205} = 0,028, \quad \sigma_{a_1} = \sqrt{0,028} = 0,167.$$

Тоді довірчі інтервали для α_0 та α_1 мають вигляд:

$$-4,26 - 2,306 \cdot 0,476 < \alpha_0 < -4,26 + 2,306 \cdot 0,476,$$

$$3,19 - 2,306 \cdot 0,167 < \alpha_1 < 3,19 + 2,306 \cdot 0,167$$

або

$$-5,36 < \alpha_0 < -3,16,$$

$$2,81 < \alpha_1 < 3,58.$$

Щоб відобразити ці довірчі інтервали графічно, в оціночне рівняння замість a_0 підставимо нижню і верхню межу для α_0 . Отримаємо рівняння: $y = -5,36 + 3,195x$ та $y = -3,16 + 3,195x$. Побудуємо ці прямі та оціночну пряму, обчисливши для кожної з них координати двох точок:

Нижня межа $y = 3,195x - 5,36$	
x	y
2	1,03
5	10,62

Оціночна пряма $y = 3,195x - 4,26$	
x	y
2	2,13
5	11,71

Верхня межа $y = 3,195x - 3,16$	
x	y
2	3,23
5	12,82

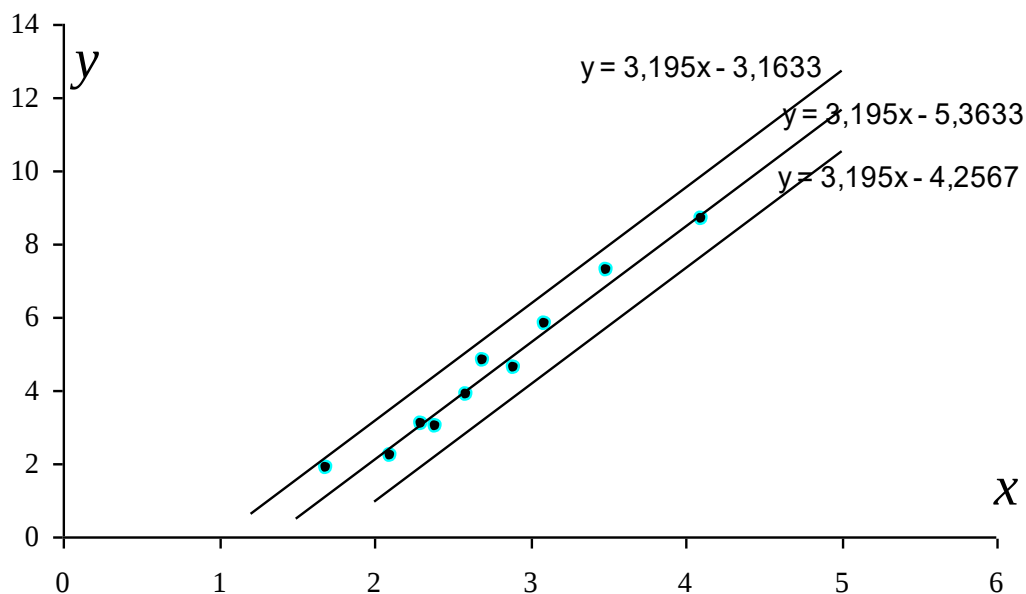


Рис. 3. Довірчий інтервал для α_0

Для побудови довірчого інтервалу для α_1 , підставимо в оціночне рівняння нижню і верхню межу для α_1 , маємо $y = -4,26 + 2,81x$ та

$y = -4,26 + 3,58x$. Обчислимо для кожної з цих прямих координати двох точок для їх побудови:

Нижня межа $y = -4,26 + 2,81x$	
x	y
0	-4,26
5	9,79

Оціночна пряма $y = -4,26 + 3,195x$	
x	y
0	-4,26
5	11,71

Верхня межа $y = -4,26 + 3,58x$	
x	y
0	-4,26
5	13,64

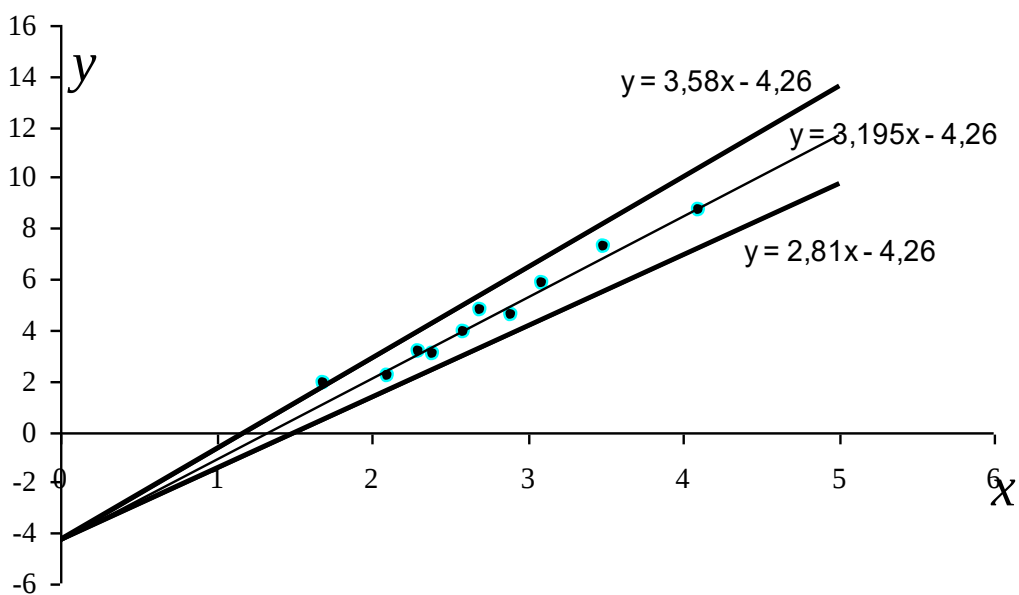


Рис. 4. Довірчий інтервал для α_1

7. Для побудови нижньої межі довірчого інтервалу економетричної моделі потрібно відкласти на координатній площині точки з координатами $\{x_j; \hat{y}_j - t(\gamma, n-2)S_{\hat{y}_j}\}$, $j = \overline{1, n}$ і з'єднати сусідні (по індексу j) точки відрізками. Щоб отримати верхню межу, робимо те ж саме з точками $\{x_j; \hat{y}_j + t(\gamma, n-2)S_{\hat{y}_j}\}$.

$$\text{Величину } S_{\hat{y}_j} \text{ знаходимо за формулою } S_{\hat{y}_j} = \sigma_u \sqrt{\left[\frac{1}{n} + \frac{(x_j - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right]}.$$

Для цього спочатку знайдемо значення $\sigma_u = \sqrt{\sigma_u^2} = \sqrt{0,1195} = 0,35$ та $\sum_{i=1}^n (x_i - \bar{x})^2 = 4,205$.

Тоді

$$S_{\hat{y}_1} = 0,35 \sqrt{\left[\frac{1}{10} + \frac{(1,8 - 2,75)^2}{4,205} \right]} = 0,196;$$

$$S_{\hat{y}_2} = 0,35 \sqrt{\left[\frac{1}{10} + \frac{(2,1 - 2,75)^2}{4,205} \right]} = 0,157;$$

$$S_{\hat{y}_3} = 0,35 \sqrt{\left[\frac{1}{10} + \frac{(2,4 - 2,75)^2}{4,205} \right]} = 0,126;$$

$$S_{\hat{y}_4} = 0,135; \quad S_{\hat{y}_5} = 0,114; \quad S_{\hat{y}_6} = 0,114;$$

$$S_{\hat{y}_7} = 0,111; \quad S_{\hat{y}_8} = 0,126; \quad S_{\hat{y}_9} = 0,169; \quad S_{\hat{y}_{10}} = 0,256.$$

З таблиці критичних значень функції Ст'юдента знаходимо $t(0,95; 8) = 2,306$ і обчислюємо ординати точок нижньої межі довірчого інтервалу:

$$\hat{y}_1 - tS_{\hat{y}_1} = 1,49 - 2,306 \cdot 0,196 = 1,037;$$

$$\hat{y}_2 - tS_{\hat{y}_2} = 2,45 - 2,306 \cdot 0,157 = 2,089;$$

$$\hat{y}_3 - tS_{\hat{y}_3} = 3,41 - 2,306 \cdot 0,126 = 3,12;$$

$$\hat{y}_4 - tS_{\hat{y}_4} = 3,09 - 2,306 \cdot 0,135 = 2,779;$$

$$\hat{y}_5 - tS_{\hat{y}_5} = 4,05 - 2,306 \cdot 0,114 = 3,778;$$

$$\hat{y}_6 - tS_{\hat{y}_6} = 5,01 - 2,306 \cdot 0,114 = 4,748;$$

$$\hat{y}_7 - tS_{\hat{y}_7} = 4,37 - 2,306 \cdot 0,111 = 4,114;$$

$$\hat{y}_8 - tS_{\hat{y}_8} = 5,65 - 2,306 \cdot 0,126 = 5,36;$$

$$\hat{y}_9 - tS_{\hat{y}_9} = 6,93 - 2,306 \cdot 0,169 = 6,54;$$

$$\hat{y}_{10} - tS_{\hat{y}_{10}} = 8,84 - 2,306 \cdot 0,256 = 8,25.$$

Тоді обчислимо значення ординат точок верхньої межі довірчого інтервалу:

$$\hat{y}_1 + tS_{\hat{y}_1} = 1,49 + 2,306 \cdot 0,196 = 1,943;$$

$$\hat{y}_2 + tS_{\hat{y}_2} = 2,45 + 2,306 \cdot 0,157 = 2,811;$$

$$\hat{y}_3 + tS_{\hat{y}_3} = 3,41 + 2,306 \cdot 0,126 = 3,7;$$

$$\hat{y}_4 + tS_{\hat{y}_4} = 3,09 + 2,306 \cdot 0,135 = 3,401;$$

$$\hat{y}_5 + tS_{\hat{y}_5} = 4,05 + 2,306 \cdot 0,114 = 4,312;$$

$$\hat{y}_6 + tS_{\hat{y}_6} = 5,01 + 2,306 \cdot 0,114 = 5,272;$$

$$\hat{y}_7 + tS_{\hat{y}_7} = 4,37 + 2,306 \cdot 0,111 = 4,626;$$

$$\hat{y}_8 + tS_{\hat{y}_8} = 5,65 + 2,306 \cdot 0,126 = 5,94;$$

$$\hat{y}_9 + tS_{\hat{y}_9} = 6,93 + 2,306 \cdot 0,169 = 7,32;$$

$$\hat{y}_{10} + tS_{\hat{y}_{10}} = 8,84 + 2,306 \cdot 0,256 = 9,43.$$

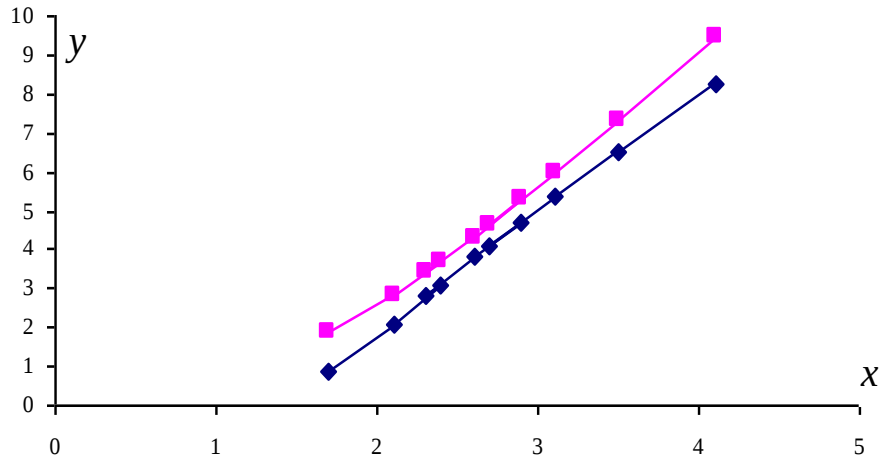


Рис. 5. Довірчий інтервал економетричної моделі

Довірчий інтервал економетричної моделі матиме вигляд як на рис. 5

8. Висуваємо нульову гіпотезу $H_0: R_{ген}=0$ (робимо припущення, що коефіцієнт кореляції генеральної сукупності рівний нулю).

Альтернативною гіпотезою буде $H_1: R_{ген} \neq 0$.

Далі для заданої вибірки обчислимо емпіричне значення параметру t :

$$t_{\text{емп}} = \frac{r\sqrt{n-m-1}}{\sqrt{1-r^2}} = \frac{0,99\sqrt{10-1-1}}{\sqrt{1-0,99^2}} = 19,8.$$

Для рівня значимості $\alpha = 0,05$ та числа ступенів вільності $k=n-2$ ($k=8$) для двосторонньої критичної області $t_{кр.} = t_{\text{довост}}(0,05;8)$ табличне значення $t_{кр.} = 2,306$.

Оскільки $|t_{\text{емп}}| > t_{кр.}$, ($19,8 > 2,306$), то з надійністю $p=0,95$ гіпотезу H_0 необхідно відхилити і прийняти альтернативну гіпотезу H_1 про існування залежності між змінними. Отже, у 95 % вибірок із генеральної сукупності коефіцієнт кореляції не дорівнює нулю.

Далі виконаємо перевірку нульової гіпотези стосовно α_1 ($H_0: \alpha_1=0$) проти альтернативної $H_1: \alpha_1 \neq 0$. Для цього знаходимо емпіричне значення t за формулою:

$$t_{\text{емп}_{\alpha_1}} = \frac{|a_1|}{s_{a_1}} = \frac{3,195}{0,167} = 19,13$$

Оскільки емпіричне значення t більше критичного ($19,3 > 2,306$), то нульова гіпотеза відхиляється і робиться висновок, що кутовий коефіцієнт a_1 , розрахований за даною вибіркою є статистично значущим з ймовірністю $p=0,95$.

Перевіримо нульову гіпотезу $H_0: \alpha_0=0$. Обчислимо

$$t_{\text{емп}_{\alpha_0}} = \frac{|a_0|}{s_{a_0}} = \frac{4,26}{0,476} = 8,95.$$

$t_{\text{емп}} > t_{кр.}$, ($19,3 > 2,306$), значить нульова гіпотеза стосовно параметру α_0 теж відхиляється, а значить α_0 не може бути рівним нулю в генеральній сукупності.

9. Для оцінки рівня адекватності побудованої економетричної моделі експериментальним даним використовуємо критерій Фішера F .

Обчислимо:

$$F_{\text{емп}} = \frac{r^2(n-m-1)}{m(1-r^2)} = \frac{0,98 \cdot (10-1-1)}{1 \cdot (1-0,98)} = 392$$

Знайдемо табличне значення даного критерію ($F_{кр.}$) для рівня надійності $p=0,95$ та числа ступенів вільності $k_1=m=1$, $k_2=n-m-1=10-1-1=8$: $F_{кр.} = 5,32$.

Оскільки $F_{\text{емп.}} > F_{кр.}$, то отримане нами оціночне рівняння економетричної моделі

$$\hat{y} = 3,195x - 4,26$$

відповідає реальній дійсності і на його основі можна здійснювати

прогнози.

Запитання для самоконтролю та самостійної роботи

1. Яка модель відноситься до категорії економетричних?
2. Яким чином визначається набір змінних для побудови економетричної моделі?
3. З яких причин у модель фактичних даних вводиться випадкова складова (залишки)?
4. Що називають специфікацією економетричної моделі?
5. У чому сутність методу найменших квадратів (МНК)?
6. Яка економічна інтерпретація параметрів рівняння регресії?
7. За допомогою яких характеристик перевіряють тісноту зв'язку між змінними моделі та значимість цього зв'язку?
8. Що виражає коефіцієнт кореляції?
9. Яка економічна інтерпретація оцінок рівняння регресії:
 - стандартної похибки оцінки за рівнянням регресії;
 - коефіцієнта детермінації;
 - кореляційного відношення;
 - вибірових похибок параметрів регресії;
 - коефіцієнта кореляції?
10. У чому суть нульової гіпотези для коефіцієнта кореляції?
11. Що показує рівень значущості в критерії Фішера?

Завдання 2. Нелінійна регресія

У таблиці наведено дані щодо випуску промислової продукції за 7 місяців. Підібрати криві, які найповніше описують тенденцію.

Дата	Випуск продукції, млрд. крб.
січень	46516
лютий	52928
березень	57928
квітень	58827
травень	59978
червень	65169
липень	64513

Використаємо можливості EXCEL. Будуємо графік даних (вибираємо тип діаграми **Точкова**, вид точкової діаграми 1). Можна забрати рамку і фон. Отримуємо діаграму розсіювання (рис. 6):

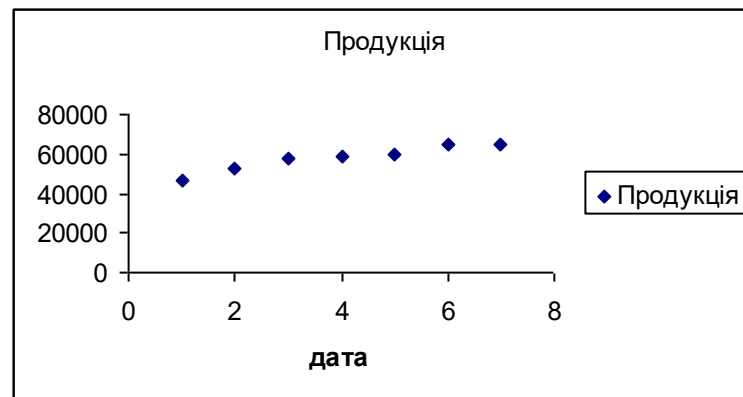


Рис.6. Діаграма розсіювання

Дальше правою кнопкою миші натискаємо по одній з точок діаграми і вибираємо з списку, що появився **Додати лінію тренду**, ЛК. У вікні, що появилось (рис. 7) вибираємо тип лінії тренду, а в **Параметрах** відмічаємо **Показати рівняння на діаграмі** і **Помістити на діаграму величину R^2** . **ОК**.

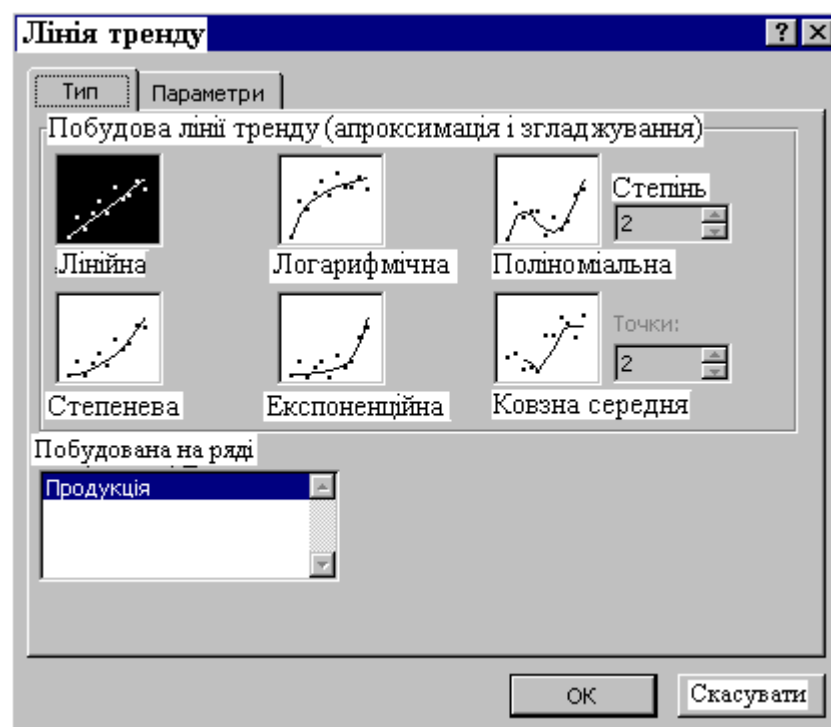


Рис. 7. Діалогове вікно **Лінія тренду**

Для кращого підбору лінії тренду пробуємо декілька варіантів типу трендів і отримуємо наступні результати:

- лінійна – $y = 2875,8x + 46477$; $r^2 = 0,9045$;
- поліноміальна (ступінь 2) – $y = -403,35x^2 + 6102,6x + 41636$; $r^2 = 0,9579$;
- поліноміальна (ступінь 3) – $y = 102,94x^3 - 1638,7x^2 + 10323x + 37930$; $r^2 = 0,9668$;
- логарифмічна – $y = 9391,3 \ln x + 46542$; $r^2 = 0,9693$;
- степенева – $y = 46899x^{0,1694}$; $r^2 = 0,9729$;
- експоненційна – $y = 46981e^{0,0511x}$; $r^2 = 0,8821$.

Найкраще апроксимує статистичні дані модель $\hat{y} = 46899x^{0,1694}$, з коефіцієнтом детермінації $r^2 = 0,9729$ (рис. 8).

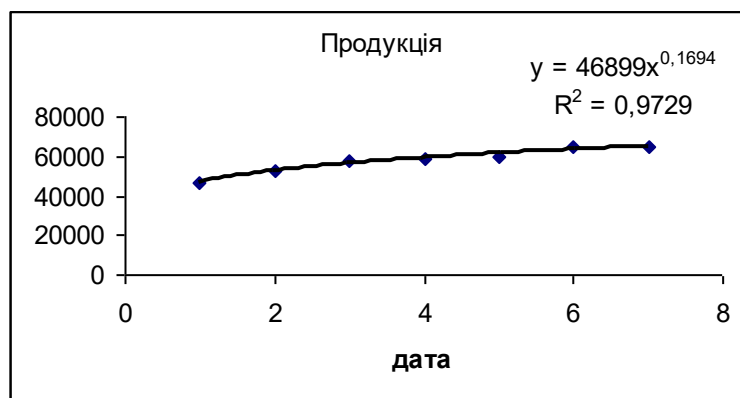


Рис. 8. Графік кривої зростання $\hat{y} = 46899x^{0,1694}$

Задача. У таблиці наведена статистична інформація для показників Y та X . Використати зворотну та степеневу модель для дослідження

залежності собівартості Y (гр.од./шт.) від кількості виготовленої продукції X (шт.).

x_i	3	4	6	7	7	9
y_i	1	2	3	3	4	6

Розв'язування.

а) Зворотна залежність має вигляд: $y = \alpha_0 + \alpha_1 \frac{1}{x} + u$.

Необхідно оцінити параметри цієї моделі. Відповідна економетрична модель має вигляд: $\hat{y} = a_0 + a_1 \frac{1}{x}$. Для лінеаризації зробимо заміну $\frac{1}{x} = t$.

Одержимо: $\hat{y} = a_0 + a_1 t$. Ця функція є лінійною. a_0 та a_1 обчислимо з системи нормальних рівнянь:

$$\begin{cases} na_0 + a_1 \sum_{i=1}^n t_i = \sum_{i=1}^n y_i, \\ a_0 \sum_{i=1}^n t_i + a_1 \sum_{i=1}^n t_i^2 = \sum_{i=1}^n y_i t_i. \end{cases}$$

Для знаходження коефіцієнтів цієї системи складемо розрахункову таблицю

Таблиця 15

i	x_i	y_i	$t_i = 1/x_i$	t_i^2	$y_i t_i$
1	1	3	1	1	3
2	2	4	0,5	0,25	2
3	3	6	0,3333	0,1111	2
4	3	7	0,3333	0,1111	2,3333
5	4	7	0,25	0,0625	1,75
6	6	9	0,1667	0,0278	1,5000
Σ		36	2,5833	1,5625	12,5833

$$\begin{cases} 6a_0 + 2,5833a_1 = 36 \\ 2,5833a_0 + 1,5625a_1 = 12,5833 \end{cases}$$

Розв'язок цієї системи:

$$a_0 = 8,789$$

$$a_1 = -6,478.$$

$$\hat{y} = 8,789 - 6,478t.$$

Отже, зворотна залежність має вигляд $\hat{y} = 8,789 - 6,478 \frac{1}{x}$.

б) Степенева залежність має вигляд: $y = \alpha_0 x^{\alpha_1}$.

Необхідно оцінити параметри цієї моделі. Відповідна економетрична модель має вигляд: $\hat{y} = a_0 x^{\alpha_1}$. Для лінеаризації логарифмуємо праву і ліву частину залежності: $\ln \hat{y} = \ln a_0 + a_1 \ln x$.

Зробимо заміну: $\ln \hat{y} = \hat{y}'$; $\ln a_0 = a'_0$; $\ln x = x'$.

Одержимо: $\hat{y}' = a'_0 + a_1 x'$

Ця функція є лінійною відносно нових змінних $\hat{y}'$ і x' .

Статистичні оцінки a'_0 , a_1 степеневого рівняння регресії із врахуванням заміни, задовольняють систему нормальних рівнянь

$$\begin{cases} na'_0 + a_1 \sum_{i=1}^n x'_i = \sum_{i=1}^n y'_i, \\ a'_0 \sum_{i=1}^n x'_i + a_1 \sum_{i=1}^n x'^2_i = \sum_{i=1}^n y'_i x'_i. \end{cases}$$

Для знаходження коефіцієнтів цієї системи складемо розрахункову таблицю

Таблиця 16

i	x_i	y_i	$x'_i = \ln x_i$	$y'_i = \ln y_i$	x'^2_i	$y'_i x'_i$
1	1	3	0	1,09861303	0	0
2	2	4	0,69314765	1,38629529	0,4804537	0,96090732
3	3	6	1,09861303	1,79176067	1,20695059	1,96845162
4	3	7	1,09861303	1,94591146	1,20695059	2,13780368
5	4	7	1,38629529	1,94591146	1,92181464	2,69760790
6	6	9	1,79176067	2,19722606	3,21040632	3,93690324
Σ			6,0684297	10,3657180	8,02657579	11,7016738

$$\begin{cases} 6a'_0 + 6,0684297 a_1 = 10,3657180 \\ 6,0684297 a'_0 + 8,02657579 a_1 = 11,7016738 \end{cases}$$

Розв'язок цієї системи:

$$a'_0 = 1,075599 \qquad a_1 = 0,644667 .$$

$$\hat{y}' = 1,075599 + 0,644667 x' .$$

Перейдемо знову до нелінійної степеневі моделі з допомогою оберненого перетворення: $a_0 = e^{a'_0} = 2,9317$.

$$\text{Отже, } \hat{y} = 2,9317x^{0,6447} .$$

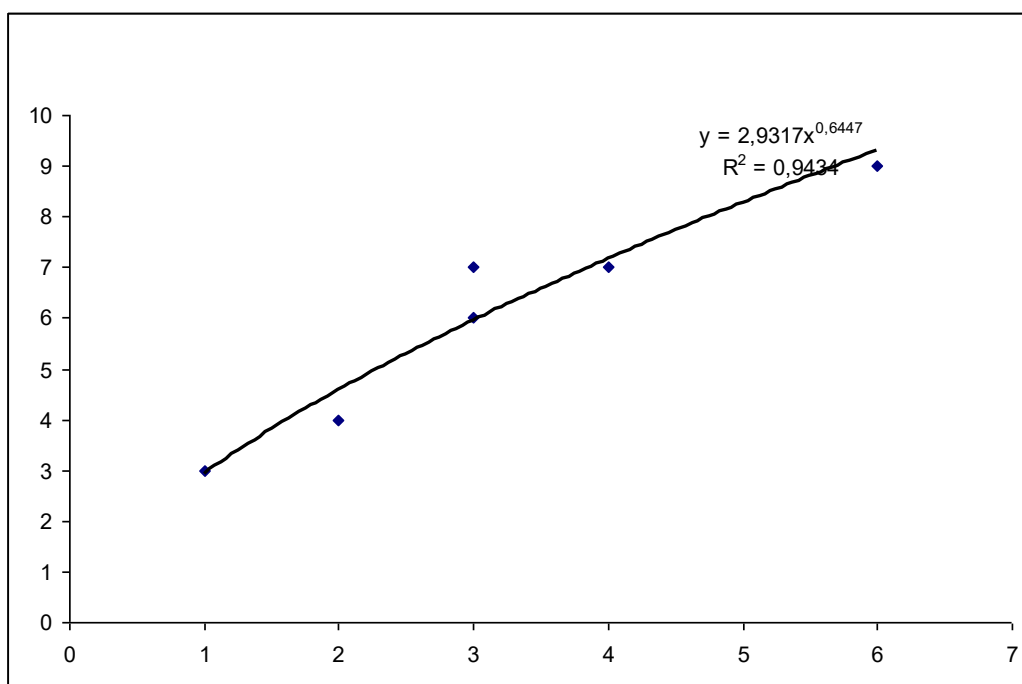


Рис. 9.

Запитання для самоконтролю та самостійної роботи

1. Дайте тлумачення нелінійної регресії та назвіть її основні види.
2. Які типи нелінійних кривих (*кривих зростання*) зустрічаються в макро- та мікроекономічних дослідженнях?
3. Якими способами лінеаризується функція відносно нових невідомих?
4. Запишіть рівняння парної та множинної регресії при параболічній залежності між змінними, поясніть їх склад.
5. Запишіть рівняння парної та множинної регресії при гіперболічній залежності між змінними, поясніть їх склад.
6. Запишіть рівняння парної та множинної регресії при степеневій залежності між змінними, поясніть їх склад.
7. Показати на графіках розповсюджені в економіці види залежностей між змінними?

Завдання 3. Багатофакторна регресія.

Знаходження оцінок методом найменших квадратів з застосуванням системи нормальних рівнянь

Задача. Задана вибірка трьох змінних Y, X_1, X_2 у вигляді таблиці

Табл. 17

Y_i	4	5	6	8	11	11	12	12	13	14
X_1	3	4	5	7	9	11	10	12	11	12
X_2	7	9	11	12	12	15	18	21	22	24

- 1) Визначити оцінки a_0, a_1, a_2 з допомогою МНК, вважаючи, що економетрична модель лінійна. Знайти ці ж оцінки, застосовуючи формули відхилення від середніх.
- 2) Обчислити коефіцієнти парної та частинної кореляції;
- 3) Знайти коефіцієнти множинної детермінації та кореляції.

1) Розв'язування. a_0, a_1, a_2 знаходимо розв'язавши систему нормальних рівнянь МНК.

$$\begin{cases} \sum_{i=1}^n Y_i = na_0 + a_1 \sum_{i=1}^n X_{1i} + a_2 \sum_{i=1}^n X_{2i} \\ \sum_{i=1}^n Y_i X_{1i} = a_0 \sum_{i=1}^n X_{1i} + a_1 \sum_{i=1}^n X_{1i}^2 + a_2 \sum_{i=1}^n X_{2i} X_{1i} \\ \sum_{i=1}^n Y_i X_{2i} = a_0 \sum_{i=1}^n X_{2i} + a_1 \sum_{i=1}^n X_{1i} X_{2i} + a_2 \sum_{i=1}^n X_{2i}^2 \end{cases}$$

Як і в попередніх роботах розрахунки сум, які входять у систему нормальних рівнянь, зводимо в таблицю.

Табл. 18

№	Y_i	X_{1i}	X_{2i}	X_{1i}^2	X_{2i}^2	$Y_i X_{1i}$	$Y_i X_{2i}$	$X_{1i} X_{2i}$
1	4	3	7	9	49	12	28	21
2	5	4	9	16	81	20	45	36
3	6	5	11	25	121	30	66	55
4	8	7	12	49	144	56	96	84
5	11	9	12	81	144	99	132	108
6	11	11	15	121	225	121	165	165
7	12	10	18	100	324	120	216	180
8	12	12	21	144	441	144	252	252
9	13	11	22	121	484	143	286	242
10	14	12	24	144	576	168	336	288
Σ	96	84	151	810	2589	913	1622	1431

Випишемо систему нормальних рівнянь, підставляючи розраховані суми

$$\begin{cases} 96 = 10a_0 + 84a_1 + 151a_2 \\ 913 = 84a_0 + 810a_1 + 1431a_2 \\ 1622 = 151a_0 + 1431a_1 + 2589a_2 \end{cases}$$

Розв'язуємо систему рівнянь з допомогою матричного методу.

Матриця коефіцієнтів системи нормальних рівнянь

$$\begin{pmatrix} 10 & 84 & 151 \\ 84 & 810 & 1431 \\ 151 & 1431 & 2589 \end{pmatrix}$$

Обернена до неї матриця

$$\begin{pmatrix} 0.848978 & -0.02401 & -0.03625 \\ -0.02401 & 0.053163 & -0.02798 \\ -0.03625 & -0.02798 & 0.017968 \end{pmatrix}$$

Розв'язки системи нормальних рівнянь, отримані множенням оберненої матриці на вектор вільних членів

$$\begin{pmatrix} 0.848978 & -0.02401 & -0.03625 \\ -0.02401 & 0.053163 & -0.02798 \\ -0.03625 & -0.02798 & 0.017968 \end{pmatrix} \times \begin{pmatrix} 96 \\ 913 \\ 1622 \end{pmatrix} = \begin{pmatrix} 0.79 \\ 0.84 \\ 0.11 \end{pmatrix}$$

Отже, розв'язок системи нормальних рівнянь буде таким:

$$a_0 = 0.79$$

$$a_1 = 0.84$$

$$a_2 = 0.11$$

Таким чином, рівняння оціночної площини набуває виду

$$\hat{Y} = 0,79 + 0,84X_1 + 0,11X_2.$$

Оцінки економетричної моделі можна розрахувати за формулами відхилення від середніх. Формули відхилення від середніх отримуємо виконанням нескладних алгебраїчних перетворень системи нормальних рівнянь.

$$a_1 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(X_{1i} - \bar{X}_1) \sum_{i=1}^n (X_{2i} - \bar{X}_2)^2 - \sum_{i=1}^n (Y_i - \bar{Y})(X_{2i} - \bar{X}_2) \sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{\sum_{i=1}^n (X_{1i} - \bar{X}_1)^2 \sum_{i=1}^n (X_{2i} - \bar{X}_2)^2 - \left(\sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) \right)^2}$$

$$a_2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(X_{2i} - \bar{X}_2) \sum_{i=1}^n (X_{1i} - \bar{X}_1)^2 - \sum_{i=1}^n (Y_i - \bar{Y})(X_{1i} - \bar{X}_1) \sum_{i=1}^n (X_{2i} - \bar{X}_2)(X_{1i} - \bar{X}_1)}{\sum_{i=1}^n (X_{2i} - \bar{X}_2)^2 \sum_{i=1}^n (X_{1i} - \bar{X}_1)^2 - \left(\sum_{i=1}^n (X_{2i} - \bar{X}_2)(X_{1i} - \bar{X}_1) \right)^2}$$

$$a_0 = \bar{Y} - \hat{\beta}_1 \bar{X}_1 - \hat{\beta}_2 \bar{X}_2$$

Обчислення заносимо в таблицю

Табл. 19

Y_i	X_{1i}	X_{2i}	$\bar{Y}$	$\bar{X}_1$	$\bar{X}_2$	$Y_i - \bar{Y}$	$X_{1i} - \bar{X}_1$	$X_{2i} - \bar{X}_2$	$(X_{1i} - \bar{X}_1)^2$	$(X_{2i} - \bar{X}_2)^2$	$(X_{1i} - \bar{X}_1) \times (X_{2i} - \bar{X}_2)$	$(Y_i - \bar{Y}) \times (X_{1i} - \bar{X}_1)$	$(Y_i - \bar{Y}) \times (X_{2i} - \bar{X}_2)$
4	3	7	9,6	8,4	15,1	-5,6	-5,4	-8,1	29,16	65,61	43,74	30,24	45,36
5	4	9				-4,6	-4,4	-6,1	19,36	37,21	26,84	20,24	28,06
6	5	11				-3,6	-3,4	-4,1	11,56	16,81	13,94	12,24	14,76
8	7	12				-1,6	-1,4	-3,1	1,96	9,61	4,34	2,24	4,96
11	9	12				1,4	0,6	-3,1	0,36	9,61	-1,86	0,84	-4,34
11	11	15				1,4	2,6	-0,1	6,76	0,01	-0,26	3,64	-0,14
12	10	18				2,4	1,6	2,9	2,56	8,41	4,64	3,84	6,96
12	12	21				2,4	3,6	5,9	12,96	34,81	21,24	8,64	14,16
13	11	22				3,4	2,6	6,9	6,76	47,61	17,94	8,84	23,46
14	12	24				4,4	3,6	8,9	12,96	79,21	32,04	15,84	39,16
96	84	151				0	0	0	104,4	308,9	162,6	106,6	172,4

$$\sum_{i=1}^{10} (Y_i - \bar{Y})^2 = (-5,6)^2 + (-4,6)^2 + (-3,6)^2 + (-1,6)^2 + (1,4)^2 + (1,4)^2 + (2,4)^2 + (2,4)^2 + (3,4)^2 + (4,4)^2 = 31,36 + 21,16 + 12,96 + 2,56 + 1,96 + 1,96 + 5,76 + 5,76 + 11,56 + 19,36 = 114,4$$

Знаходимо оцінки економетричної моделі

$$a_1 = \frac{106,6 \cdot 308,9 - 172,4 \cdot 162,6}{104,4 \cdot 308,9 - (162,6)^2} \approx 0,8427$$

$$a_2 = \frac{172,4 \cdot 104,4 - 106,6 \cdot 162,6}{104,4 \cdot 308,9 - (162,6)^2} \approx 0,1145$$

$$a_0 = 9,6 - 0,84 \cdot 8,4 - 0,1145 \cdot 15,1 = 0,7919$$

Отже, $\hat{Y} = 0,79 + 0,84X_1 + 0,11X_2$.

Результати розрахованих двома способами оцінок повинні співпасти.

2) Перш, ніж досліджувати економетричну модель з трьома змінними $\hat{Y} = a_0 + a_1X_1 + a_2X_2$, необхідно розрахувати три парних моделі наступного виду

$$Y = \alpha' + \beta'X_1$$

$$Y = \alpha'' + \beta''X_2$$

$$X_1 = \alpha''' + \beta'''X_2$$

Для кожної моделі розраховуємо коефіцієнти кореляції, які називають коефіцієнтами кореляції нульового порядку чи коефіцієнтами парної кореляції:

$$r_{YX_1} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(X_{1i} - \bar{X}_1)}{\sqrt{\sum (Y_i - \bar{Y})^2 \sum (X_{1i} - \bar{X}_1)^2}}, \quad r_{YX_2} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(X_{2i} - \bar{X}_2)}{\sqrt{\sum (Y_i - \bar{Y})^2 \sum (X_{2i} - \bar{X}_2)^2}},$$

$$r_{X_1X_2} = \frac{\sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2)}{\sqrt{\sum (X_{2i} - \bar{X}_2)^2 \sum (X_{1i} - \bar{X}_1)^2}}$$

Для розрахунку даних коефіцієнтів скористаємося обчисленими в попередньому пункті задачі значеннями із таблиці 19.

$$r_{YX_1} = \frac{106.4}{\sqrt{114.4 \cdot 104.4}} = \frac{106.4}{109.28} \approx 0.97, \quad r_{YX_2} = \frac{172.4}{\sqrt{114.4 \cdot 308.9}} = \frac{172.4}{187.98} \approx 0.92,$$

$$r_{X_1X_2} = \frac{162.6}{\sqrt{104.4 \cdot 308.9}} = \frac{162.6}{179.58} \approx 0.91$$

Матриця коефіцієнтів кореляції нульового порядку набуває вигляду

$$|r| = \begin{pmatrix} 1 & 0.97 & 0.92 \\ 0.97 & 1 & 0.91 \\ 0.92 & 0.91 & 1 \end{pmatrix}.$$

Із допомогою матриці коефіцієнтів парної кореляції можна здійснити аналіз дисперсій змінних моделі. Проведення такого аналізу є одним із основних завдань економетрії. Значення коефіцієнта кореляції $r_{YX_1}=0.97$, $r_{YX_1}^2 \approx 0.94$. Це означає, що 94% загальної дисперсії змінної Y пояснюються економетричною моделлю $Y = \alpha' + \beta'X_1$. Непояснена дисперсія, викликана випадковою складовою u , складає 6%.

Коефіцієнт $r_{YX_2}=0.92$, $r_{YX_2}^2 \approx 0.85$. Це означає, що 85% загальної дисперсії Y пояснюються економетричною моделлю $Y = \alpha'' + \beta''X_2$. Непояснена дисперсія складає 15%.

Очевидно, що якщо обмежитися розрахунками тільки перших моделей, то слід вибрати першу модель, яка пояснює більшу частку загальної дисперсії змінної Y .

Окремо слід зупинитися на аналізі дисперсій пояснюючих змінних X_1, X_2 який базується на коефіцієнті кореляції $r_{X_1X_2}$.

Значення $r_{X_1X_2} = 0.91$, $r_{X_1X_2}^2 = 0.83$ показує, що 83% загальної дисперсії змінної X_1 пояснюється економетричною моделлю $X_1 = \alpha''' + \beta'''X_2$.

Іншими словами, пояснююча змінна X_1 корелює з пояснюючою змінною X_2 . Велике значення коефіцієнта кореляції між пояснюючими змінними економетричної моделі свідчить про явище мультиколінеарності, яке полягає в

тому, що вплив пояснюючої змінної X_1 на результативну Y можна вважати опосередкованим через змінну X_2 і навпаки.

Здійснивши аналіз дисперсій на основі коефіцієнтів парної кореляції, можна поставити запитання в іншому аспекті. А саме, якщо $r_{YX_1}^2 \approx 0,94$, то непоясненою залишається 6% дисперсії. Якщо в модель ввести ще одну пояснюючу змінну X_2 , то необхідно дізнатися яку частку дисперсії непоясненої змінною X_1 пояснить введення змінної X_2 . Визначити цю частку можна з допомогою коефіцієнта частинної кореляції чи коефіцієнта кореляції першого порядку

$$r_{YX_2 \cdot X_1} = \frac{r_{YX_2} - r_{YX_1} \cdot r_{X_1X_2}}{\sqrt{1 - r_{YX_1}^2} \cdot \sqrt{1 - r_{X_1X_2}^2}} = \frac{0.92 - 0.97 \cdot 0.91}{\sqrt{1 - (0.97)^2} \cdot \sqrt{1 - (0.91)^2}} = 0.4.$$

Отже, якщо $r_{YX_2 \cdot X_1} = 0.4$, то з 6% дисперсії змінної Y , непоясненої змінною X_1 40% пояснює введення ще однієї пояснюючої змінної X_2 .

Коефіцієнт частинної кореляції $r_{YX_1 \cdot X_2}$ визначається наступним чином

$$r_{YX_1 \cdot X_2} = \frac{r_{YX_1} - r_{YX_2} \cdot r_{X_1X_2}}{\sqrt{1 - r_{YX_2}^2} \cdot \sqrt{1 - r_{X_1X_2}^2}} = \frac{0.97 - 0.92 \cdot 0.91}{\sqrt{1 - (0.92)^2} \cdot \sqrt{1 - (0.91)^2}} = 0.83.$$

Значить, із 15% дисперсії змінної Y непоясненої змінною X_2 , 83% пояснює введення в модель змінної X_1 .

Отже, аналіз дисперсій, який базується на матриці коефіцієнтів парної кореляції та коефіцієнтів частинної кореляції дає важливу інформацію при розрахунку економетричної моделі

$$\hat{Y} = a_0 + a_1 X_1 + a_2 X_2.$$

Висновок.

1) Модель виду $\hat{Y} = a_0 + a_1 X_1 + a_2 X_2$ можна замінити моделлю виду $Y = \alpha' + \beta' X_1$ на підставі того, що ця модель пояснює 94% дисперсії змінної Y . Введення в модель додаткової змінної X_2 пояснює лише 40% дисперсії Y , непоясненої змінною X_1 . Змінна X_2 призводить до ускладнення моделі та збільшення обсягу розрахунків.

2) Модель виду $\hat{Y} = a_0 + a_1 X_1 + a_2 X_2$ не варто замінювати моделлю виду $Y = \alpha'' + \beta'' X_2$ оскільки ця модель пояснює лише 85% дисперсії змінної Y . Введення в модель додаткової змінної X_1 пояснює лише 83% дисперсії Y , непоясненої змінною X_2 .

Тепер визначаємо коефіцієнт множинної детермінації, який показує частку загальної дисперсії змінної Y , що пояснює оціночна площина.

$$R^2 = \frac{r_{YX_1}^2 + r_{YX_2}^2 - 2r_{YX_1}r_{YX_2}r_{X_1X_2}}{1 - r_{X_1X_2}^2}$$

$$R^2 = \frac{(0.97)^2 + (0.92)^2 - 2 \cdot 0.97 \cdot 0.92 \cdot 0.91}{1 - (0.91)^2} = 0.95$$

Коефіцієнт множинної кореляції або коефіцієнт другого порядку

$$R = \sqrt{R^2} = \sqrt{0.95} = 0.97$$

Запитання для самоконтролю та самостійної роботи

1. У чому різниця між парним і множинним рівнянням регресії?
2. Яка економічна інтерпретація параметрів множинного рівняння регресії?
3. Які ви знаєте способи подання системи нормальних рівнянь у багатовимірному способі?
4. Побудуйте систему нормальних рівнянь для економетричної моделі в якій є три незалежних змінних.
5. Сформулюйте припущення методу найменших квадратів.
6. Дайте тлумачення коефіцієнтів парної кореляції та запишіть формули їхнього застосування.
7. Як визначають матрицю коефіцієнтів парної кореляції ?
8. Що відображає коефіцієнт частинної кореляції?
9. Дайте тлумачення коефіцієнтів частинної кореляції та запишіть формули для їх знаходження.
10. Чим відрізняються коефіцієнти парної та частинної кореляції?
11. Як визначається коефіцієнт множинної кореляції?
12. Як визначається коефіцієнт множинної детермінації та його зв'язок із оціненим коефіцієнтом детермінації?

Завдання 4. Багатофакторна регресія.

Знаходження оцінок методом найменших квадратів з застосуванням матричної форми запису

Використовуючи вибірккові дані попереднього завдання, знайти:

1. Вектор оцінок $\mathbf{a} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$

2. Матрицю дисперсій оцінок $\text{var}(\mathbf{a})$

3. Інтервали довір'я ($p=0,9$) параметрів α_0 , α_1 та α_2 .

Допустимо, що між показником y і факторами x_1 , x_2 існує лінійна залежність

$$\hat{y} = a_1 + a_2 x_1 + a_3 x_2$$

Знайдемо оцінки параметрів, використовуючи матричні операції за формулою

$$\mathbf{a} = \left[[X]^T [X] \right]^{-1} [X]^T Y, \text{ де}$$

$$[X] = \begin{pmatrix} 1 & 3 & 7 \\ 1 & 4 & 9 \\ 1 & 5 & 11 \\ 1 & 7 & 12 \\ 1 & 9 & 12 \\ 1 & 11 & 15 \\ 1 & 10 & 18 \\ 1 & 12 & 21 \\ 1 & 11 & 22 \\ 1 & 12 & 24 \end{pmatrix} - \text{матриця пояснюючих змінних } X_1, X_2, \text{ доповнена колонкою}$$

$$\text{одиниць, } Y = \begin{pmatrix} 4 \\ 5 \\ 6 \\ 8 \\ 11 \\ 11 \\ 12 \\ 12 \\ 13 \\ 14 \end{pmatrix} - \text{вектор результативної змінної } Y, [X]^T - \text{транспонована}$$

матриця пояснюючих змінних.

Проведемо відповідно обчислення

$$[X]^T [X] = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 3 & 4 & 5 & 7 & 9 & 11 & 10 & 12 & 11 & 12 \\ 7 & 9 & 11 & 12 & 12 & 15 & 18 & 21 & 22 & 24 \end{pmatrix} \times \begin{pmatrix} 1 & 3 & 7 \\ 1 & 4 & 9 \\ 1 & 5 & 11 \\ 1 & 7 & 12 \\ 1 & 9 & 12 \\ 1 & 11 & 15 \\ 1 & 10 & 18 \\ 1 & 12 & 21 \\ 1 & 11 & 22 \\ 1 & 12 & 24 \end{pmatrix} =$$

$$= \begin{pmatrix} 10 & 84 & 151 \\ 84 & 810 & 1431 \\ 151 & 1431 & 2589 \end{pmatrix}.$$

$$[[X]^T [X]]^{-1} = \begin{pmatrix} 0,848978 & -0,02401 & -0,03625 \\ -0,02401 & 0,053163 & -0,02798 \\ -0,03625 & -0,02798 & 0,017968 \end{pmatrix}.$$

Наступним кроком є обчислення добутку матриці $[X]^T$ і вектора Y :

$$[X]^T Y = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 3 & 4 & 5 & 7 & 9 & 11 & 10 & 12 & 11 & 12 \\ 7 & 9 & 11 & 12 & 12 & 15 & 18 & 21 & 22 & 24 \end{pmatrix} \times \begin{pmatrix} 4 \\ 5 \\ 6 \\ 8 \\ 11 \\ 11 \\ 12 \\ 12 \\ 13 \\ 14 \end{pmatrix} = \begin{pmatrix} 96 \\ 913 \\ 1622 \end{pmatrix}.$$

Тоді вектор оцінок

$$\mathbf{a} = \begin{pmatrix} 0,848978 & -0,02401 & -0,03625 \\ -0,02401 & 0,053163 & -0,02798 \\ -0,03625 & -0,02798 & 0,017968 \end{pmatrix} \times \begin{pmatrix} 96 \\ 913 \\ 1622 \end{pmatrix} = \begin{pmatrix} 0,79 \\ 0,84 \\ 0,11 \end{pmatrix}.$$

Отже економетрична модель має вигляд:

$$y = 0,79 + 0,84x_1 + 0,11x_2.$$

Знаходимо матрицю дисперсій оцінок :

$$\text{VAR}(\mathbf{a}) = \sigma_u^2 [[X]^T [X]]^{-1} = \begin{pmatrix} \sigma_{a_0}^2 & \sigma_{a_0 a_1}^2 & \sigma_{a_0 a_2}^2 \\ \sigma_{a_1 a_0}^2 & \sigma_{a_1}^2 & \sigma_{a_1 a_2}^2 \\ \sigma_{a_2 a_0}^2 & \sigma_{a_2 a_1}^2 & \sigma_{a_2}^2 \end{pmatrix},$$

де $\sigma_u^2 = \frac{u'u}{n-k}$, u – вектор відхилень значень змінної Y_i із вибірки від

розрахункових значень цієї змінної $\hat{Y}_i$,

u' - транспонований вектор до, n – об'єм вибірки, k – кількість змінних, які входять в модель.

Діагональні елементи матриці визначають дисперсії оцінок a_0 , a_1 , a_2 . Інші

елементи взаємну коваріацію цих оцінок. З допомогою діагональних елементів матриці $\text{VAR}(\mathbf{a})$ знаходяться граничні похибки оцінок для заданого рівня ймовірності.

Знаходимо розрахункові значення $\hat{Y}$

$$\hat{y}_1 = 0,79 + 0,84 \cdot 3 + 0,11 \cdot 7 = 4,12; \quad \hat{y}_2 = 0,79 + 0,84 \cdot 4 + 0,11 \cdot 9 = 5,19; \quad \text{і т. д.}$$

$$\hat{y}_{10} = 0,79 + 0,84 \cdot 12 + 0,11 \cdot 24 = 13,65.$$

Обчислюємо вектор відхилень:

$$u_1 = 4 - 4,12 = -0,12; \quad u_2 = 5 - 5,19 = -0,19; \quad \text{і т. д.} \quad u_{10} = 14 - 13,65 = 0,35.$$

$$u'u =$$

$$\begin{pmatrix} -0,12 & -0,19 & -0,26 & -0,06 & 1,25 & -0,78 & 0,72 & -1,31 & 0,42 & 0,35 \end{pmatrix} \cdot \begin{pmatrix} -0,12 \\ -0,19 \\ -0,26 \\ -0,06 \\ 1,25 \\ -0,78 \\ 0,72 \\ -1,31 \\ 0,42 \\ 0,35 \end{pmatrix} =$$

$$= 4,82.$$

$$\sigma_u^2 = 4,82 / (10 - 3) = 0,69.$$

$$\begin{aligned} \text{var}(\mathbf{a}) &= 0,69 \times \begin{pmatrix} 0,848978 & -0,02401 & -0,03625 \\ -0,02401 & 0,053163 & -0,02798 \\ -0,03625 & -0,02798 & 0,017968 \end{pmatrix} = \\ &= \begin{pmatrix} 0,585795 & -0,01657 & -0,02501 \\ -0,01657 & 0,036682 & -0,01931 \\ -0,02501 & -0,01931 & 0,012398 \end{pmatrix}. \end{aligned}$$

Запитання для самоконтролю та самостійної роботи

1. Опишіть основні припущення лінійної багатофакторної моделі в матричній формі.
2. Опишіть етапи алгоритму визначення параметрів багатофакторної моделі.
3. Яка структура дисперсійно-коваріаційної матриці та для чого її розраховують?
4. Як визначити довірчі інтервали параметрів моделі?

Завдання 5. Мультиколінеарність в економетричних моделях

Економічний показник y залежить від трьох факторів x_1, x_2, x_3 . На основі статистичних даних за 10 років необхідно: оцінити наявність загальної мультиколінеарності. У випадку її присутності, виявити пари факторів між якими існує мультиколінеарність, один із факторів виключити з розгляду таких пар.

Табл. 20

x_1	x_2	x_3
1,43	9,18	8,97
3,92	10,94	9,8
5,49	11,3	9,6
8,17	13,99	8,02
9,68	14,59	10,51
14,42	15,73	12,39
15,92	17,97	16,36
18,04	19,45	9,32
20,69	21,49	12,72
22,68	21,8	16,22

Для знаходження кореляційної матриці використовуємо вбудовану функцію **CORREL** (коефіцієнт парної кореляції), яка знаходиться в категорії **СТАТИСТИКА** або заходимо в **Вставлення – Аналіз Даних – Кореляція**.

$$[R] = \begin{bmatrix} 1 & 0,99 & 0,70 \\ 0,99 & 1 & 0,66 \\ 0,70 & 0,66 & 1 \end{bmatrix}. \quad \text{Для знаходження } \chi_p^2 \text{ знаходимо}$$

визначник матриці $[R]$ $\det[R] = 0,0077$. Визначник зручно обчислити з допомогою вбудованої математичної функції **MDETERM** (блок кореляційної матриці). Розрахункове значення χ_i^2 знаходимо за формулою

$$\chi_p^2 = -[n - 1 - \frac{1}{6}(2m + 5)] \ln \det[R] = -\left[10 - 1 - \frac{1}{6}(6 + 5)\right] \ln 0,0077 = 34,85.$$

Для довірчої ймовірності $p=0,95$ і числа ступенів вільності $\frac{1}{2}m(m-1) = \frac{3}{2}(3-1) = 3$, $\chi_{кр}^2(0,95;3) = 7,8$.

Так як розрахункове значення $\chi_p^2 = 34,85$ і воно більше критичного, то з надійністю 0,95 можна стверджувати, що існує загальна мультиколінеарність. Пари факторів між якими існує мультиколінеарність знаходимо використовуючи t -статистику. Знаходимо обернену матрицю до кореляційної:

$$[Z] = [R]^{-1} = \begin{bmatrix} 72,48 & -67,98 & -5,83 \\ -67,98 & 65,54 & 4,29 \\ -5,83 & 4,29 & 2,25 \end{bmatrix}$$

t - статистику пари факторів розрахуємо за формулою:

$$t_{ij} = \frac{r_{ij}^* \sqrt{n-m-1}}{\sqrt{1-r_{ij}^{*2}}},$$

де $r_{ij}^* = \frac{-z_{ij}}{\sqrt{z_{ii}z_{jj}}}$; z_{ij}, z_{ii}, z_{jj} – елементи матриці $[Z]$.

$$r_{12}^* = \frac{-(-67,98)}{\sqrt{72,48 \cdot 65,54}} = 0,986; \quad t_{12} = \frac{0,986 \cdot \sqrt{10-3-1}}{\sqrt{1-0,986^2}} = 14,639;$$

$$r_{13}^* = \frac{-(-5,83)}{\sqrt{72,48 \cdot 2,25}} = 0,456; \quad t_{13} = \frac{0,456 \cdot \sqrt{10-3-1}}{\sqrt{1-0,456^2}} = 1,256;$$

$$r_{23}^* = \frac{-4,29}{\sqrt{65,54 \cdot 2,25}} = -0,353; \quad t_{23} = \frac{-0,353 \cdot \sqrt{10-3-1}}{\sqrt{1-0,353^2}} = -0,923.$$

Для ступенів вільності $k=n-m-1=10-3-1=6$ і $p=0,95$ критичне значення $t(0,95;6) = 2,227$. Отже, звідси видно, що лише для пари факторів x_1 і x_2 $t_{12} > t(0,95;6)$, тобто з надійністю $p=0,95$ між факторами x_1 і x_2 існує мультиколінеарність. Виключаємо з розгляду один із факторів, нехай це буде x_1 .

Наступним кроком буде знаходження кореляційної матриці між факторами x_2 і x_3 : $[R] = \begin{bmatrix} 1 & 0,66 \\ 0,66 & 1 \end{bmatrix}$.

Обернена $[Z] = \begin{bmatrix} 1,785 & -1,184 \\ -1,184 & 1,785 \end{bmatrix}$, визначник матриці кореляції $\det[R] = 0,560$.

Значення $\chi_p^2 = -\left[10-1-\frac{1}{6}(4+5)\right] \ln 0,560 = 4,346$ менше від критичного $\chi_{кр}^2(0,95;2) = 6,0$, це значить, що загальна мультиколінеарність між факторами x_2, x_3 – відсутня.

Запитання для самоконтролю та самостійної роботи

1. Що означає мультиколінеарність змінних?
2. Дайте коротку характеристику алгоритму Феррара-Глобера.
3. Які існують способи усунення мультиколінеарності?

Завдання №6

ТЕСТИ ДЛЯ ВИЯВЛЕННЯ ОЗНАКИ ГЕТЕРОСКЕДАСТИЧНОСТІ

Тест рангової кореляції Спірмена. Дана процедура є найбільш простою та доступною серед множини аналітичних методів. Її можна використовувати як для малих, так і для великих вибірок. Основу алгоритму даного тесту складає обчислення коефіцієнта рангової кореляції Спірмена:

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n^3 - n}, \quad (6.1)$$

де d_i – різниця між рангами, які властиві двом характеристикам i -го об'єкта; n – кількість об'єктів, що рангуються.

Алгоритм тесту рангової кореляції Спірмена складається із таких кроків.

Припустимо, що $y_i = \beta_0 + \beta_1 x_i + u_i$.

Крок 1. Будуємо рівняння регресії $\hat{Y} = b_0 + b_1 x$.

Крок 2. Знаходимо відхилення u_i .

Крок 3. Нехтуючи знаком u_i , тобто розглядаємо $|u_i|$, ранжуємо їх та x_i у зростаючому чи спадному порядку.

Крок 4. Знаходимо d_i – різницю між рангами, а також d_i^2 .

Крок 5. Знаходимо за формулою (6.1) коефіцієнт r_s .

Крок 6. Перевіряємо значущість отриманого коефіцієнта рангової кореляції за t -критерієм Ст'юдента. Розрахункове (емпіричне) значення t -критерію знаходимо за формулою:

$$t_{emn} = \frac{r_s \sqrt{n-2}}{\sqrt{1-r_s^2}}, \quad (6.2)$$

де n – кількість спостережень; $k = n - 2$ – число ступенів вільності.

Крок 7. Знаходимо критичне значення t -критерію, тобто $t_{кр}$.

Крок 8. Якщо $t_{emn} > t_{кр}$, то підтверджується гіпотеза про гетероскедастичність. Якщо дана нерівність не виконується, то має місце гомоскедастичність.

Приклад

За даними нижче приведеної таблиці про витрати обігу (Y) та вантажообіг (X) побудована економетрична модель $\hat{Y} = 29,46 - 0,74X$. На рівні значущості $\alpha = 0,05$ перевірити наявність гетероскедастичності в економетричній моделі, використовуючи тести: рангової кореляції Спірмена, Гольдфельда–Квандта, Глейзера.

Розв’язання. Стосовно нашого прикладу:

Кроки 1, 2. Побудуємо модель $\hat{Y} = 29,46 - 0,74X$ та знайдемо відхилення $u_i = y_i - \hat{y}_i$

Крок 3. Здійснимо ранжування змінних $|u_i|$ та x_i . При заповненні сьомого стовпця вибирається найменше число в п’ятому стовпці і поруч з ним записується 1. Найменшому з тих чисел, що залишилися, відповідає 2, і т.д.

Крок 4. Знаходимо $d_i = \text{ранг}|u_i| - \text{ранг} x_i$, а також d_i^2 .

Розрахунки проведемо в таблиці

Табл. 21

Витрати обігу, тис. грн. Y	Вантажообіг тис. грн, X	$\hat{Y}$	$Y - \hat{Y}$ u	$ Y - \hat{Y} $ $ u $	Ранг	Ранг $ u $	d	d^2
1	2	3	4	5	6	7	8	9
25	6	25,02	-0,02	0,02	1	2	1	1
24	10	22,06	1,94	1,94	2	10	8	64
20	12	20,58	-0,58	0,58	3	4	1	1
20	14	19,1	0,9	0,9	4	6	2	4
17	15	18,36	-1,36	1,36	5	7	2	4
14	20	14,66	-0,66	0,66	6	5	-1	1
10	24	11,7	-1,7	1,7	7	9	2	4
9	27	9,48	-0,48	0,48	8	3	-5	25
8	29	8	0	0	9	1	-8	64
5	35	3,56	1,44	1,44	10	8	-2	4
								172

Крок 5. Обчислимо коефіцієнт рангової кореляції Спірмена за формулою (6.1).

$$r_s = 1 - \frac{6 \cdot 172}{10^3 - 10} = 1 - \frac{1032}{990} \approx -0,04242.$$

Крок 6. Перевіримо значущість отриманого коефіцієнта рангової кореляції за t -критерієм Ст'юдента. Розрахункове значення t -критерію знаходимо за формулою (6.2)

$$t_{emp} = \frac{|-0,04242| \cdot \sqrt{10-2}}{\sqrt{1-(-0,04242)^2}} \approx 0,12009.$$

Крок 7. Знаходимо критичне значення t -критерію для $k=10-2=8$ числа ступенів вільності й рівня значущості $\alpha = 0,05$ ($t_{кр}=2,306$).

Крок 8. Оскільки $0,12009 < 2,306$, то гіпотеза про гетероскедастичність не підтверджується.

Параметричний тест Гольдфельда-Квандта. Даний метод використовується, коли вхідна сукупність спостережень невелика. В його основу покладено припущення відносно того, що дисперсія залишків зростає пропорційно до величини квадрата однієї з незалежних змінних, тобто має місце формула $M(u^2) = \sigma_u^2 X_{ij}^2$ для побудованої нами моделі виду

$$Y = XB + U \quad (6.3)$$

Зауваження 1. Перевірка за даним тестом повинна здійснюватися окремо для кожного регресора (незалежної змінної), що входить у багатофакторну модель.

Алгоритм даного тесту включає в себе наступні кроки:

1. Впорядковуємо вхідну сукупність спостережень відповідно до величини зростання елементів вектора X_j , який найбільш імовірно може викликати зміну дисперсії залишків.

2. Відкидаємо із вхідної сукупності C спостережень, які містяться в центрі вектора X_j . Величину C знаходимо за формулою:

$$C = \frac{4n}{15}, \quad (6.5)$$

де n – кількість елементів вектора X_j . У результаті такої процедури залишок $(n-C)$ спостережень поділиться на дві підвибірки однакового розміру $(n-C)/2$, одна з яких буде включати малі значення вектора X_j , а інша – великі.

3. Використовуючи МНК (систему нормальних рівнянь або формули відхилення від середнього) або стандартну процедуру продукту EXCEL, побудуємо економетричні моделі для кожної з

отриманих на попередньому кроці підвибірок, обсяг яких становить $\frac{n-C}{2} \geq m$, де m – кількість незалежних змінних. Як результат одержимо дві моделі виду:

$$Y_1 = XB_1 + U_1, \quad Y_2 = XB_2 + U_2. \quad (6.5)$$

Зауваження 2. Якщо $n_1 \neq n_2$, то відкидається перше або останнє спостереження сукупності.

4. Знаходимо суму квадратів залишків для першої та другої моделі:

$$S_1 = U_1^T U_1, \quad S_2 = U_2^T U_2, \quad (6.6)$$

де S_1, S_2 – сума квадратів залишків, відповідно для першої та другої моделі; U_1, U_2 – вектори залишків, відповідно для першої та другої моделі.

5. Обчислюємо величину критерію:

$$F^* = \frac{S_2}{S_1}. \quad (6.7)$$

6. Знаходимо критичне значення критерію Фішера ($F_{кр}$) для $k_1 = k_2 = \frac{n-C}{2} - m - 1$ ступенів вільності та рівня значущості $\alpha = 0,05$.

7. Якщо $F^* > F_{кр}$ ми допускаємо існування гетероскедастичності (відкидаємо гіпотезу H_0 щодо відсутності відмінності між дисперсіями випадкових величин U в двох підвибірках). Якщо $F^* \leq F_{кр}$, то гетероскедастичність відсутня (приймається гіпотеза H_0). Чим більше значення F^* , тим більша гетероскедастичність залишків.

Покажемо застосування даного тесту для нашої задачі.

Крок 1. Упорядкуємо вхідну сукупність спостережень відповідно до величини зростання елементів вектора X_1 , який найбільш імовірно може викликати зміну дисперсії залишків.

Зауваження 3. Для цього можна використати процедуру сортування засобів ПК.

Y	X ₁
25	6
24	10
20	12
20	14
17	15
14	20
10	24
9	27
8	29
5	35

Крок 2. Розрахуємо кількість спостережень, які слід відкинути із середини сукупності

$$C = \frac{4 \cdot 10}{15} \approx 2,7$$

У нашому випадку викидаємо 5-е та 6-е спостереження. Тоді одержану вибірку ділимо на дві підвибірки однакового розміру $\frac{10-2}{2} = 4$.

Крок 3. Оціночне рівняння економетричних моделей кожної підвибірки можна розрахувати в EXCEL, користуючись функцією “РЕГРЕСІЯ” або статистичною функцією “LINEST” (див. додаток).

Для нашого прикладу

Лінійн1:

-0,7	29,6
0,226779	2,473863
0,826506	1,341641
9,527778	2
17,15	3,6

Лінійн2:

-0,46332	21,32046
0,027842	0,80826
0,99283	0,224038
276,9231	2
13,89961	0,100386

Таким чином, оціночне рівняння для першої підвибірки має вигляд $\hat{y}^1 = 29,6 - 0,7X$ для другої – $\hat{y}^2 = 21,32 - 0,46X$

Крок 4. Знайдемо суму квадратів залишків для першої та другої моделі. Ці суми містяться у п'ятому рядку результатів функції “LINEST” (виділені сірим кольором). Отже,

$$S_1=3,6, \quad S_2=0,1.$$

Крок 5. Обчислюємо величину критерію F^* за формулою (6.7)

$$F^* = \frac{0,1}{3,6} = 0,027778.$$

Крок 6. Знаходимо табличне значення критерію Фішера для $k_1 = k_2 = \frac{n-C}{2} - m - 1 = \frac{10-2}{2} - 1 - 1 = 2$ ступенів вільності і рівня значущості $\alpha = 0,05$ ($F_{кр} = 19$).

Зауваження 4. Число ступенів вільності вказане також у четвертому рядку результатів функції "LINEST".

Крок 7. Оскільки $0,027778 < 19$, то гетероскедастичність відсутня, що підтверджує достовірність побудованої моделі в завданні 1.

Тест Глейзера. Базується на визначенні регресійної залежності між модулем залишків та тією змінною, яка може спричинитися до гетероскедастичності, тобто розглядається регресія

$$|u_i| = f(x_i) + \varepsilon_i, \quad i = \overline{1, n}. \quad (6.8)$$

В якості функції f зазвичай обирається функція виду

$$f(x) = \beta_0 + \beta_1 x^\delta, \quad (6.9)$$

Регресія (6.8) вивчається при різних значеннях δ , а потім вибирається те конкретне значення, при якому коефіцієнт β_1 виявляється найбільш значущим, тобто має найбільше значення t -статистики. При цьому в якості значень δ беруться числа: 1, 2, 3, 1/2, 1/3 тощо. Якщо ж β_1 незначущий для всіх розглянутих значень δ (випадок $f \equiv \text{const}$), тоді робиться висновок про відсутність гетероскедастичності.

Крок 1. Використовуючи звичайний МНК, визначаємо за даною вибіркою статистичні оцінки a та b у випадку однофакторної економетричної моделі або b_0 та b_j ($j=1, m$) у випадку багатофакторної економетричної моделі.

Крок 2. Записуємо оціночне рівняння моделі. Визначаємо відхилення u_i .

Крок 3. Будуємо модель залежності відхилень $|u_i|$ від обраного регресора x_i у вигляді (6.8).

Крок 4. Для вибраного значення δ лінеаризованої моделі, використовуючи звичайний МНК, визначають статистичні оцінки b_0 , b_1 та S_{b_j} .

Крок 5. Здійснюється перевірка на статистичну значущість всіх параметрів моделі при заданому рівні значущості α , використовуючи критерій Ст'юдента (t -критерій):

$$t_{\beta_j} = \frac{|b_j|}{S_{b_j}}, \quad (6.10)$$

де b_j – оцінка відповідного параметра моделі; $S_{b_j} = \sqrt{S_{b_j}^2}$,

$S_{b_j}^2$ – незміщена оцінка невідомої дисперсії параметра β_j

Крок 6. Знаходимо критичне значення ($t_{кр}$) критерію Ст'юдента для $k = n - m - 1$ ступенів вільності та рівня значущості α .

Крок 7. Порівнюємо розраховані значення t_{β_j} з критичним. У випадку, якщо $t_{\beta_j} \leq t_{кр}$, оцінки параметрів є статистично незначущі, а це означає, що залишки не є гетероскедастичними. Якщо ж $t_{\beta_j} > t_{кр}$, тоді приймаємо гіпотезу про наявність гетероскедастичності залишків.

Стосовно нашого прикладу:

Крок 1, 2. Нами побудована економетрична модель $\hat{Y} = 29,46 - 0,74X$ та знайдені відхилення $u_i = y_i - \hat{y}_i$

x_i	y_i	$\hat{y}_i$	$y_i - \hat{y}_i$	$ u_i = y_i - \hat{y}_i $
6	25	25,02	-0,02	0,02
10	24	22,06	1,94	1,94
12	20	20,58	-0,58	0,58
14	20	19,1	0,9	0,9
15	17	18,36	-1,36	1,36
20	14	14,66	-0,66	0,66
24	10	11,7	-1,7	1,7
27	9	9,48	-0,48	0,48
29	8	8	0	0
35	5	3,56	1,44	1,44

Крок 3. Покладемо у функції (6.9) $\delta = 1$, тобто розглядатимемо модель $|u_i| = \beta_0 + \beta_1 x$.

Крок 4. Використовуючи функцію “LINEST” знайдемо статистичні оцінки b_0 b_1 та S_{b_j} .

0,0031874	0,8468024
0,0256564	0,5425588
0,0019255	0,7191122
0,0154338	8
0,0079812	4,1369788

Таким чином, $|u_i| = 0,8468 + 0,0032x$,

$$S_{b_0} = 0,5426, \quad S_{b_1} = 0,0257$$

Крок 5. Визначаємо достовірність параметрів моделі за формулою (6.10):

$$t_{\beta_0} = \frac{0,8468024}{0,5425588} = 1,560757$$

$$t_{\beta_1} = \frac{0,0031874}{0,0256564} = 0,124233$$

Крок 6. Для рівня значущості 0,05 і числа ступенів вільності $k = 10 - 1 - 1 = 8$ критичне значення критерію Ст’юдента $t_{кр} = 2,306$.

Крок 7. Оскільки $t_{\beta_j} < t_{кр}$ ($1,560757 < 2,306$; $0,124233 < 2,306$), гіпотезу про наявність гетероскедастичності відкидаємо.

Завдання №7 АВТОКОРЕЛЯЦІЯ

Критерій Дарбіна-Уотсона. При використанні даного критерію формулюється нульова гіпотеза про відсутність автокореляції $H_0 : \rho = 0$. Альтернативна гіпотеза може бути побудованою на основі використання односторонньої критичної області – $H_1 : \rho > 0$, тобто існує додатна, або $H_1 : \rho < 0$, тобто існує від’ємна автокореляція, або на основі двосторонньої критичної області – $H_1 : \rho \neq 0$.

Статистика Дарбіна-Уотсона визначається за формулою:

$$DW = d = \frac{\sum_{t=2}^n (u_t - u_{t-1})^2}{\sum_{t=1}^n u_t^2} \quad (7.1)$$

Між статистикою d і коефіцієнтом авторегресії ρ існує наближене співвідношення:

$$d \approx 2 - 2\rho. \quad (7.2)$$

Оскільки $-1 < \rho < 1$, то для загального випадку має місце $0 < d < 4$.

Розподіл статистики d (7.2), як випадкової величини, залежить від кількості спостережень у вибірці n , числа пояснюючих змінних у рівнянні регресії m , конкретних значень пояснюючих змінних і рівня значущості α . Ці обставини свідчать про неможливість побудови таблиці критичних значень d -статистики. Тому Дарбіном і Уотсоном було доведено, що існує нижня та верхня границі для критичного значення d , які не залежать від значень вибірки, а визначаються тільки обсягом вибірки n , числом пояснюючих змінних m і рівнем значущості ε .

Позначимо нижню границю величини d через d_n , а верхню – через d_ε . Тому в таблиці d -статистики приведені значення d_n та d_ε для рівнів значущості $\alpha = 0,01$ ($\alpha = 0,025$, $\alpha = 0,05$) при n від 15 до 100 і числа пояснюючих змінних m , від 1 до 5.

Зона автокореляційного зв'язку за критерієм Дарбіна-Уотсона

Відкидається H_0 Приймається H_1	?	Приймається H_0 про відсутність автокореляції залишків		?	Відкидається H_0 Приймається H_1
0 d_n	d_6	2	$4-d_6$	$4-d_n$	4
Додатна автокореляція	Зона невизна ченості	Автокореляція відсутня		Зона невизна ченості	Від’ємна автокореляція

Критерій Дарбіна-Уотсона має два недоліки. Перший з них полягає в тому, що критичні границі прийняття гіпотези H_0 і відхилення альтернативної H_1 не співпадають. Критичні значення утворюють п'ять областей різних статистичних рішень, у тому числі області невизначеності. В областях невизначеності нульова гіпотеза не приймається і не відкидається, тобто з допомогою d -критерію неможливо прийти до якогось висновку. Другий недолік полягає в тому, що при обсязі вибірки $n < 15$ для d не існує критичних значень d_n і d_e .

Знайдене значення d -статистики порівнюється з табличними значеннями d_n і d_e , і, керуючись наступними правилами, робляться висновки:

1. $d_e \leq d \leq 4-d_e$ – приймається гіпотеза $H_0 : \rho = 0$, тобто автокореляція відсутня;
2. $0 < d \leq d_n$ – приймається гіпотеза $H_1 : \rho > 0$, тобто існує додатна автокореляція залишків;
3. $d_n < d < d_e$ та $4-d_e < d < 4-d_n$ – при вибраному рівні значущості неможливо прийти до певного висновку (виникає потреба в додатковому дослідженні);
4. $4-d_n \leq d < 4$ – приймається гіпотеза $H_1 : \rho < 0$, тобто має місце від'ємна автокореляція.

Якщо залишки не містять автокореляцію, тобто $\rho=0$, то значення d міститься поблизу числа 2. Дане твердження доцільно використовувати в тому випадку, коли потрібно впевнитися у тому, є автокореляція чи ні.

Приклад

На основі даних про витрати обігу (Y) та вантажообіг (X) на рівні значущості $\alpha = 0,05$ перевірити наявність автокореляції в економетричній моделі, використовуючи критерій Дарбіна–Уотсона.

Розв’язання: Крок. 1. За даними нижче приведеної таблиці побудована економетрична модель $\hat{y} = 29,46 - 0,74x$ та знайдені відхилення $u_t = y_t - \hat{y}_t$

Табл. 22

x_t	y_t	$\hat{y}_t$	u_t	u_t^2	$u_t - u_{t-1}$	$(u_t - u_{t-1})^2$
6	25	25,02	-0,02	0,0004		
10	24	22,06	1,94	3,7636	1,96	3,8416
12	20	20,58	-0,58	0,3364	-2,52	6,3504
14	20	19,1	0,9	0,81	1,48	2,1904
15	17	18,36	-1,36	1,8496	-2,26	5,1076
20	14	14,66	-0,66	0,4356	0,7	0,49
24	10	11,7	-1,7	2,89	-1,04	1,0816
27	9	9,48	-0,48	0,2304	1,22	1,4884
29	8	8	0	0	0,48	0,2304
35	5	3,56	1,44	2,0736	1,44	2,0736
				12,3896		22,854

Крок 2. Знайдемо оцінку критерію Дарбіна-Уотсона за формулою (7.1):

$$DW = d = \frac{22,854}{12,3896} = 1,844612$$

Крок 3. Використовуючи табличні значення статистики Дарбіна-Уотсона для $n=10$ та $\alpha = 0,05$, одержуємо $d_n=0,879$, $d_g=1,32$.

Крок 4. Оскільки $d=1,844612$, то вона знаходиться в межах від d_g до $4-d_g$ ($1,32 < d < 2,68$). Таким чином, робимо висновок, що автокореляція відсутня.

Завдання. Лагова модель.

Задача. Розрахувати значення коефіцієнтів взаємної кореляції та побудувати графік кореляційної функції для наступних статистичних даних залежності витрат на харчування (y_t , грош. од.) від рівня доходів (x_t , грош. од.) при максимальному значенні лага часу, рівному восьми місячним періодам.

Рік	y_t	x_t
1	1400	1150
2	1560	1270
3	1720	1390
4	1500	1400
5	1250	1200
6	1320	1350
7	1429	1400
8	1560	1410
9	1600	1450
10	1713	1500
11	1800	1400
12	1780	1530
13	1850	1540
14	1900	1500
15	2100	1600

♦Розв'язування.

Розрахуємо значення кореляційної функції при $\tau = \overline{0,8}$ та відобразимо їх у наступній таблиці:

τ	0	1	2	3	4	5	6	7	8
r_τ	0,8115	0,601	0,5021	0,6359	0,7038	0,7644	0,6736	0,6837	0,6136

Побудуємо графік кореляційної функції (рис.1).

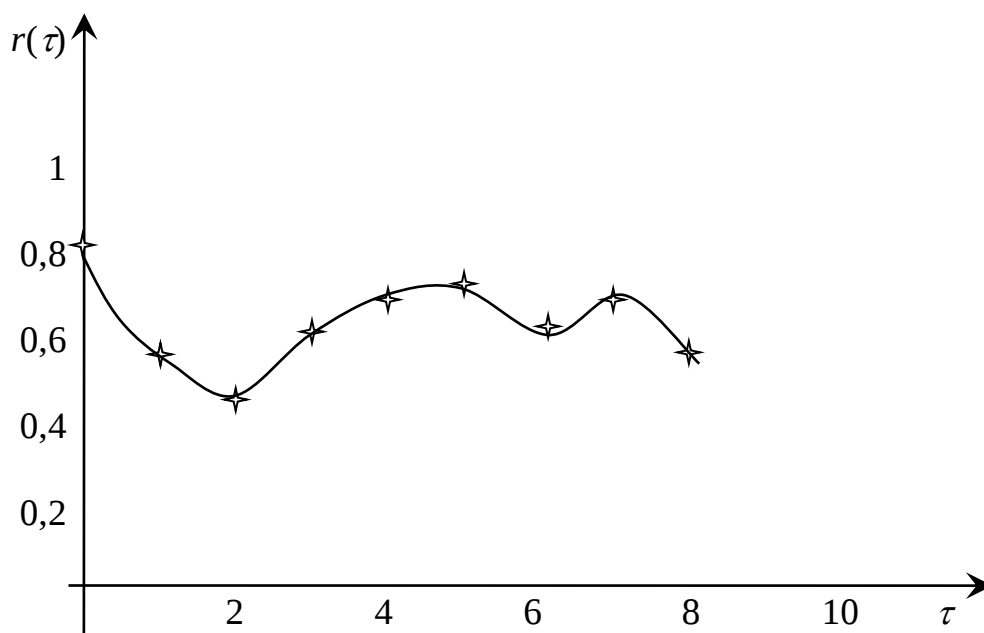


Рис. 1

Аналізуючи значення коефіцієнтів взаємної кореляційної функції, робимо висновок, що рівень доходів значно впливає на рівень витрат на харчування протягом всіх восьми періодів часу. При цьому коефіцієнт парної кореляції для показників є найбільшим ($\tau=0$), що свідчить про найбільшу схильність людей витратити кошти на харчування у найближчому поточному періоді.

ДОДАТКИ

Додаток 1

Алгоритм знаходження оберненої матриці в Excel

1. Відмічаємо поле, де буде знаходитись результат.
2. Входимо в майстер функцій (*f*). У категоріях вибираємо Математичні, а в функціях **MINVERSE**. Вводимо адресу матриці для якої знаходимо обернену. Натискаємо клавішу **Enter** або **OK**.
3. Для того щоб отримати на екрані значення всієї оберненої матриці, натискаємо клавіші **F2** і **Ctrl+Shift+Enter**.

Додаток 2

Алгоритм множення матриць в Excel

1. Відмічаємо поле, де буде знаходитись результат.
2. Входимо в майстер функцій (**f**). У категоріях вибираємо математику, а в функціях **MMULT** Вводимо адресу 1-ої матриці в 1-ий масив і 2-ої - в другий. Натискаємо клавішу **Ok**.
3. Для того щоб отримати на екрані значення всієї матриці добутку, натискаємо клавіші **F2** і **Ctrl+Shift+Enter**.

Додаток 3

Порядок знаходження оцінок параметрів економетричної моделі з використанням функції **LINEST**

1. Відмічаємо блок, де мають знаходитись розрахункові дані. Висота цього блоку завжди дорівнює 5-ти рядкам, а ширина - числу оцінюваних параметрів.
2. В діалоговому вікні **Вставлення функції**, вибираємо категорію **СТАТИСТИЧНІ** функцію **LINEST** і натискаємо на кнопку **Далі>**.
3. В наступному діалоговому вікні вводимо: в перший рядок блок даних показника; в другий рядок блок даних факторів; в третій рядок вводиться слово **TRUE**, якщо a_0 не дорівнює нулю, і слово **FALSE**, якщо a_0 дорівнює нулю; в четвертий рядок вводиться слово **TRUE**, якщо необхідно знайти не лише параметри регресії, а й додаткову регресійну статистику, і слово **FALSE** якщо необхідно знайти лише параметри регресії. Після цього натискаємо кнопку **Enter** або **OK**.
4. Для того щоб отримати на екрані значення всіх розрахункових даних, натискаємо клавіші **F2** і **Ctrl+Shift+Enter**.

Таблиця розрахункових значень має вигляд (таблиця для трьох оцінюваних параметрів):

a_2	a_1	a_0
σ_{a2}	σ_{a1}	σ_{a0}
r^2	σ_{xy}	# Н / Д
$F_{роз}$	k	# Н / Д
$\sum(\hat{y} - \bar{y})^2$	$\sum(\hat{y}_i - y_i)^2$	# Н / Д

В першому рядку знаходяться оцінки економетричної моделі відповідно a_2 , a_1 , a_0 .

В другому рядку знаходяться середні квадратичні відхилення оцінок параметрів σ_{a_2} , σ_{a_1} , σ_{a_0} .

В третьому рядку в першій комірці знаходиться коефіцієнт детермінації, а в другій комірці - середнє квадратичне відхилення показника (стандартна похибка оцінки за рівнянням економетричної моделі).

В четвертому рядку в першій комірці знаходиться розрахункове значення F - статистики, в другій комірці - k - число ступенів вільності.

В п'ятому рядку в першій комірці знаходиться сума квадратів відхилень розрахункових значень показника від його середнього значення, в другій комірці – сума квадратів відхилень розрахункових значень показника від статистичних.

Додаток 4. Довідкові таблиці

Таблиця значень функції Лапласа $\Phi(x) = \frac{1}{\pi} \int_0^x e^{-\frac{t^2}{2}} dt$

Таблиця 4.1

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0,00	0,0000	0,50	0,1915	1,00	0,3413	1,50	0,4332	2,00	0,4772	3,00	0,49865
0,01	0,0040	0,51	0,1950	1,01	0,3438	1,51	0,4345	2,02	0,4783	3,20	0,49931
0,02	0,0080	0,52	0,1985	1,02	0,3461	1,52	0,4357	2,04	0,4793	3,40	0,49966
0,03	0,0120	0,53	0,2019	1,03	0,3485	1,53	0,4370	2,06	0,4803	3,60	0,499841
0,04	0,0160	0,54	0,2054	1,04	0,3508	1,54	0,4382	2,08	0,4812	3,80	0,499928
0,05	0,0199	0,55	0,2088	1,05	0,3531	1,55	0,4394	2,10	0,4821	4,00	0,499968
0,06	0,0239	0,56	0,2123	1,06	0,3554	1,56	0,4406	2,12	0,4830	4,50	0,499997
0,07	0,0279	0,57	0,2157	1,07	0,3577	1,57	0,4418	2,14	0,4838	5,00	0,499997
0,08	0,0319	0,58	0,2190	1,08	0,3599	1,58	0,4429	2,16	0,4846		
0,09	0,0359	0,59	0,2224	1,09	0,3621	1,59	0,4441	2,18	0,4854		
0,10	0,0398	0,60	0,2257	1,10	0,3643	1,60	0,4452	2,20	0,4861		
0,11	0,0438	0,61	0,2291	1,11	0,3665	1,61	0,4463	2,22	0,4868		
0,12	0,0478	0,62	0,2324	1,12	0,3686	1,62	0,4474	2,24	0,4875		
0,13	0,0517	0,63	0,2357	1,13	0,3708	1,63	0,4484	2,26	0,4881		
0,14	0,0557	0,64	0,2389	1,14	0,3729	1,64	0,4495	2,28	0,4887		
0,15	0,0596	0,65	0,2422	1,15	0,3749	1,65	0,4505	2,30	0,4893		
0,16	0,0636	0,66	0,2454	1,16	0,3770	1,66	0,4515	2,32	0,4898		
0,17	0,0675	0,67	0,2486	1,17	0,3790	1,67	0,4525	2,34	0,4904		
0,18	0,0714	0,68	0,2517	1,18	0,3810	1,68	0,4535	2,36	0,4909		
0,19	0,0753	0,69	0,2549	1,19	0,3830	1,69	0,4545	2,38	0,4913		
0,20	0,0793	0,70	0,2580	1,20	0,3849	1,70	0,4554	2,40	0,4918		
0,21	0,0832	0,71	0,2611	1,21	0,3869	1,71	0,4564	2,42	0,4922		
0,22	0,0871	0,72	0,2642	1,22	0,3883	1,72	0,4573	2,44	0,4927		
0,23	0,0910	0,73	0,2673	1,23	0,3907	1,73	0,4582	2,46	0,4931		
0,24	0,0948	0,74	0,2703	1,24	0,3925	1,74	0,4591	2,48	0,4934		
0,25	0,0987	0,75	0,2734	1,25	0,3944	1,75	0,4599	2,50	0,4938		
0,26	0,1026	0,76	0,2764	1,26	0,3962	1,76	0,4608	2,52	0,4941		
0,27	0,1064	0,77	0,2794	1,27	0,3980	1,77	0,4616	2,54	0,4945		
0,28	0,1103	0,78	0,2823	1,28	0,3997	1,78	0,4625	2,56	0,4948		
0,29	0,1141	0,79	0,2852	1,29	0,4015	1,79	0,4633	2,58	0,4951		
0,30	0,1179	0,80	0,2881	1,30	0,4032	1,80	0,4641	2,60	0,4953		
0,31	0,1217	0,81	0,2910	1,31	0,4049	1,81	0,4649	2,62	0,4956		
0,32	0,1255	0,82	0,2939	1,32	0,4066	1,82	0,4656	2,64	0,4959		
0,33	0,1293	0,83	0,2967	1,33	0,4082	1,83	0,4664	2,66	0,4961		
0,34	0,1331	0,84	0,2995	1,34	0,4099	1,84	0,4671	2,68	0,4963		
0,35	0,1368	0,85	0,3023	1,35	0,4115	1,85	0,4678	2,70	0,4965		
0,36	0,1406	0,86	0,3051	1,36	0,4131	1,86	0,4686	2,72	0,4967		
0,37	0,1443	0,87	0,3078	1,37	0,4147	1,87	0,4693	2,74	0,4969		
0,38	0,1480	0,88	0,3106	1,38	0,4162	1,88	0,4699	2,76	0,4971		
0,39	0,1517	0,89	0,3133	1,39	0,4177	1,89	0,4706	2,78	0,4973		
0,40	0,1554	0,90	0,3159	1,40	0,4192	1,90	0,4713	2,80	0,4974		

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,41	0,1591	0,91	0,3186	1,41	0,4207	1,91	0,4719	2,82	0,4976		
0,42	0,1628	0,92	0,3212	1,42	0,4222	1,92	0,4726	2,84	0,4977		
0,43	0,1664	0,93	0,3238	1,43	0,4236	1,93	0,4732	2,86	0,4979		
0,44	0,1700	0,94	0,3264	1,44	0,4251	1,94	0,4738	2,88	0,4980		
0,45	0,1736	0,95	0,3289	1,45	0,4265	1,95	0,4744	2,90	0,4981		
0,46	0,1772	0,96	0,3315	1,46	0,4279	1,96	0,4750	2,92	0,4982		
0,47	0,1808	0,97	0,3340	1,47	0,4292	1,97	0,4756	2,94	0,4984		
0,48	0,1844	0,98	0,3365	1,48	0,4306	1,98	0,4761	2,96	0,4985		
0,49	0,1879	0,99	0,3389	1,49	0,4319	1,99	0,4767	2,98	0,4986		

Таблиця критичних точок розподілу Ст'юдента

Таблиця 4.2

Число ступенів вільності k	Рівень значимості α (двостороння критична область)					
	0,1	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,96
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29
	0,05	0,025	0,01	0,005	0,001	0,0005
	Рівень значимості α (одностороння критична область)					

Таблиця критичних точок розподілу χ^2 -квадрат

Таблиця 4.3

Число ступенів вільності k	Рівень значимості α					
	0,01	0,025	0,05	0,95	0,975	0,99
1	6,6	5,0	3,8	0,0039	0,00098	0,00016
2	9,2	7,4	6,0	0,103	0,051	0,020
3	11,3	9,4	7,8	0,352	0,216	0,115
4	13,3	11,1	9,5	0,711	0,484	0,297
5	15,1	12,8	11,1	1,15	0,831	0,554
6	16,8	14,4	12,6	1,64	1,24	0,872
7	18,5	16,0	14,1	2,17	1,69	1,24
8	20,1	17,5	15,5	2,73	2,18	1,65
9	21,7	19,0	16,9	3,33	2,70	2,09
10	23,2	20,5	18,3	3,94	3,25	2,56
11	24,7	21,9	19,7	4,57	3,82	3,05
12	26,2	23,3	21,0	5,23	4,40	3,57
13	27,7	24,7	22,4	5,89	5,01	4,11
14	29,1	26,1	23,7	6,57	5,63	4,66
15	30,6	27,5	25,0	7,26	6,26	5,23
16	32,0	28,8	26,3	7,96	6,91	5,81
17	33,4	30,2	27,6	8,67	7,56	6,41
18	34,8	31,5	28,9	9,39	8,23	7,01
19	36,2	32,9	30,1	10,1	8,91	7,63
20	37,6	34,2	31,4	10,9	9,59	8,26
21	38,9	35,5	32,7	11,6	10,3	8,90
22	40,3	36,8	33,9	12,3	11,0	9,54
23	41,6	38,1	35,2	13,1	11,7	10,2
24	43,0	39,4	36,4	13,8	12,4	10,9
25	44,3	40,6	37,7	14,6	13,1	11,5
26	45,6	41,9	38,9	15,4	13,8	12,2
27	47,0	43,2	40,1	16,2	14,6	12,9
28	48,3	44,5	41,3	16,9	15,3	13,6
29	49,6	45,7	42,6	17,7	16,0	14,3
30	50,9	47,0	43,8	18,5	16,8	15,0

Значення статистики Дарбіна-Уотсона при рівні значимості $\alpha = 0,05$

Таблиця 4.4

n	K=1		K=2		K=3		K=4		K=5	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0,61	1,40	-	-	-	-	-	-	-	-
7	0,7	1,36	0,47	1,9	-	-	-	-	-	-
8	0,76	1,33	0,56	1,78	0,37	2,29	-	-	-	-
9	0,82	1,32	0,63	1,7	0,46	2,13	-	-	-	-
10	0,88	1,32	0,7	1,64	0,53	2,02	-	-	-	-
11	0,93	1,32	0,66	1,6	0,6	1,93	-	-	-	-
12	0,97	1,33	0,81	1,58	0,66	1,86	-	-	-	-
13	1,01	1,34	0,86	1,56	0,72	1,82	-	-	-	-
14	1,05	1,35	0,91	1,55	0,77	1,78	-	-	-	-
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,90	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,30	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,88
27	1,32	1,47	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,65	1,14	1,74	0,07	1,83

Значення F-критерію Фішера при рівні значущості $\alpha=0,05$

Таблиця 4.5

$k_1 \backslash k_2$	1	2	3	4	5	6	8	12	24	∞
1	161,45	199,50	215,72	224,57	230,17	233,97	238,89	243,91	249,04	254,32
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35

80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26
125	3,92	3,07	2,68	2,44	2,29	2,17	2,01	1,83	1,60	1,21
150	3,90	3,06	2,66	2,43	2,27	2,16	2,00	1,82	1,59	1,18
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,80	1,57	1,14
300	3,87	3,03	2,64	2,41	2,25	2,13	1,97	1,79	1,55	1,10
400	3,86	3,02	2,63	2,40	2,24	2,12	1,96	1,78	1,54	1,07
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06
1000	3,85	3,00	2,61	2,38	2,22	2,10	1,95	1,76	1,53	1,03
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

СПИСОК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ

1. Єрмоєнко В., Алілуйко А., Березька К., Мартинюк О. Економетрика : навчальний посібник. Тернопіль: Підручники і посібники, 2023. 168 с.
2. Березька К. М. Тестові завдання з дисципліни «Економетрика»: навч.-метод. вказівки. Тернопіль : ЗУНУ, 2021. 95 с.
3. Березька К. М., Мартинюк О. М., Пласконь С. А., Єрмоєнко В. О., Руська Р. В., Маслій В. В. Комплексні практичні індивідуальні завдання з курсу «Економетрика». Тернопіль: ЗУНУ, 2022. 68 с.
4. Березька К.М. Економетрика: основи теорії та комп'ютерний практикум. Тернопіль: ЗУНУ, 2022. 152 с.
5. Березька К.М., Пласконь С.А., Мартинюк О.М., Єрмоєнко В.О., Дзюбановська Н.В., Руська Р.В., Маслій В.В. Методичні вказівки до виконання тренінгових завдань з дисципліни «Економетрика» для студентів денної форми навчання. Тернопіль: ЗУНУ, 2022. 30 с.
6. Диха М. В., Мороз В. С. Економетрія: Навчальний посібник. К.: Центр навчальної літератури (ЦУЛ), 2019. 206 с.
7. Іващук О. Т., Дзюбановська Н. В. Методичні рекомендації для підготовки до практичних занять з дисципліни «Економетрика». Тернопіль: ЗУНУ, 2022. 159 с.
8. Ковальчук О. Я. Математичне моделювання та прогнозування в міжнародних відносинах: Підручник. Тернопіль: ТНЕУ, 2019. 412 с.
9. Козьменко О. В., Кузьменко О. В. Економіко-математичні методи та моделі (економетрика): Навч. посібник. Суми: Університетська книга, 2018. 406 с.
10. Моделі сталого розвитку: колективна монографія / за ред. Мартинюк О.М. Вид-во Підручники і посібники. Тернопіль, 2022. 400 с.
11. Руська Р. В. Економетрика: навчальний посібник. видання 2-е перероб. доп. Тернопіль: ЗУНУ, 2022. 224 с.
12. Пласконь С., Сенів Г., Хома-Могильська С., Кармелюк Г. Економетричне оцінювання впливу валового внутрішнього продукту на рівень життя населення України. Економічний аналіз: Тернопіль, 2019. Том 29. № 3. С. 12-20.
13. Пласконь С., Сенів Г., Руська Р., Новосад І. Математико-статистичні аспекти аналізу динаміки показників заробітної плати в Україні. Економічний аналіз: Тернопіль, 2021. Том 31, № 2. С. 55-61.
14. O. Kochan, Z. Wang, Y. Ouyang, V. Eromenko, A. Aliluiko and K. Przystupa, "Criteria of Goodness of Fit and Confidence Intervals for Polynomial Regression Models Through the Origin (i.e. Without the Intercept)," 2023 14th International Conference on Measurement, Smolenice, Slovakia, 2023, pp. 43-46.
15. Berezka, K. M., Kneysler, O. V., Spasiv N. Ya., & Kulyna, H. M. (2021). Information technology for forecasting financial results of insurance companies. *Ukrainian Journal of Information Technologies*, 3(2), 87-93.
16. Gumenna-Derij, M., Khorunzhak, N., Poprozman, N., Berezka, K., Kruchak, L. (2022). Modeling, accounting and control of formation and use of

resources (on the example of the construction industry). *Independent Journal of Management & Production (Special Edition ISE, S&P)*. Special Edition ISE, S&P, Vol. 13, No. 3, 123-144.

17. Malyniak, B., Martyniuk, O., Kyrylenko, O. Corruption and efficiency of public spending in states with various public management types | Коррупция и эффективность общественных расходов стран с различными типами режимов публичного управления. *Economic Annals-XXI*[this link is disabled](#), 2019, 178(7-8), pp. 17–27.

18. Shkolnyk I., Kozmenko S., Kozmenko O., Mershchii B. The impact of the economy financialization on the level of economic development of the associate EU member states. *Economics & Sociology*, 2019. P. 43-58.

19. Stavytskyy, A., Kharlamova, G., Giedraitis, V., Cheberyako, O., & Nikytenko, D. Gender question: Econometric answer. *Economics and Sociology*, 2020. 13(4). P. 241- 255.